Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVI (2016), 845-907

Criterion of the L^2 boundedness and sharp endpoint estimates for singular integral operators on product spaces of homogeneous type

YONGSHENG HAN, JI LI AND CHIN-CHENG LIN

Abstract. The purpose of this paper is to introduce a class of general singular integral operators on spaces $\widetilde{M} = M_1 \times \cdots \times M_n$. Each factor space M_i, $1 \le i \le n$, is a space of homogeneous type in the sense of Coifman and Weiss. These operators generalize those studied by Journé on the Euclidean space and include operators studied by Nagel and Stein on Carnot-Carathéodory spaces on which the basic geometry is given by a control, or Carnot-Carathéodory, metric induced by a collection of vector fields of finite type. We provide the criterion of the $L^2(\widetilde{M})$ boundedness for these general operators. Thus this result extends the product $T1$ theorem of Journé on Euclidean space and recovers the L^p, $1 < p < \infty$, boundedness of those operators on Carnot-Carathéodory space obtained by Nagel and Stein. We also prove the sharp endpoint estimates for these general operators on the Hardy spaces $H^p(\widetilde{M})$ and $BMO(\widetilde{M})$.

Mathematics Subject Classification (2010): 42B20 (primary); 42B25 (secondary).

1. Introduction and statement of main results

Classical Calderón-Zygmund theory centers around singular integrals associated with the Hardy-Littlewood maximal operator that commutes with the usual dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta x_1, \ldots, \delta x_n)$ for $\delta > 0$. On the other hand, the product theory on $\mathbb{R}^n$ began with Zygmund's study of the strong maximal function in [16] given by

$$M_S(f)(x) = \sup_{x \in R} \frac{1}{|R|} \int_R |f(y)| dy,$$

The second author was supported by ARC DP 160100153 and Macquarie University New Staff Grant. The third author was supported by MOST of Taiwan under Grant #MOST 103-2115-M-008-003-MY3.

Received November 1, 2014; accepted in revised form April 23, 2015.
Published online September 2016.

where the supremum is taken over the family of all rectangles R with sides parallel to the axes.

And it continued with Marcinkiewicz's proof of his famous multiplier theorem. The product theory is invariant with respect to n-fold dilation on $\mathbb{R}^n$, $\delta \cdot x = (\delta_1 x_1, \dots, \delta_n x_n)$ for $\delta = (\delta_1, \dots, \delta_n) \in \mathbb{R}^n_+$. In the product setting, one considers operators of the form $Tf = K * f$, where K is homogeneous, that is, $\delta_1 \dots \delta_n K(\delta \cdot x) = K(x)$, or, more generally, $K(x)$ satisfies certain differential inequalities and cancellation conditions such that the kernels $\delta_1 \dots \delta_n K(\delta \cdot x)$ also satisfy the same conditions with the same bounds uniformly for all $\delta_i > 0, 1 \le i \le n$. Such operators have been studied for example in Gundy and Stein [11], R. Fefferman and Stein [10], R. Fefferman [8,9], Chang and R. Fefferman [2,3], Journé [17,18], Pipher [25], Pott and Villarroya [26], where both the L^p theory for $1 < p < \infty$ and the product H^p theory for $0 < p \le 1$ were developed. More precisely, R. Fefferman and Stein [10] studied the L^p boundedness $(1 < p < \infty)$ for the product convolution singular integral operators. Journé in [17] introduced a non-convolution product singular integral operators, established the product $T1$ theorem (see [26] for a new version of such operators) and proved the $L^\infty \to BMO$ boundedness for such operators. The product Hardy space $H^p(\mathbb{R}^n \times \mathbb{R}^m)$ was first introduced by Gundy and Stein [11]. Chang and R. Fefferman [2,3] developed the theory of atomic decomposition and established the dual space of Hardy space $H^1(\mathbb{R}^n \times \mathbb{R}^m)$, namely the product $BMO(\mathbb{R}^n \times \mathbb{R}^m)$ space. Carleson disproved by a counter-example a conjecture that the product atomic Hardy space on $\mathbb{R}^n \times \mathbb{R}^m$ could be defined by rectangle atoms. This motivated Chang and R. Fefferman to replace the role of cubes in the classical atomic decomposition of $H^p(\mathbb{R}^n)$ by arbitrary open sets of finite measures in the product $H^p(\mathbb{R}^n \times \mathbb{R}^m)$. Subsequently, R. Fefferman in [9] established the criterion of the $H^p \to L^p$ boundedness of singular integral operators in Journé's class by considering its action only on rectangle atoms by using Journé lemma. However, R. Fefferman's criterion cannot be extended to three or more parameters without further assumptions on the nature of T as shown in Journé [18]. In fact, Journé provided a counter-example in the three-parameter setting of singular integral operators such that R. Fefferman's criterion breaks down. The H^p to L^p boundedness for Journé's class of singular integral operators with arbitrary number of parameters was established by J. Pipher [25] by considering directly the action of the operator on (non-rectangle) atoms and an extension of Journé's geometric lemma to higher dimensions. The criterion of the $H^p \to H^p$ boundedness of singular integral operators in Journé's class on the Euclidean space was established in [12].

To study fundamental solutions of $\Box_b$ on certain model domains in several complex variables, Nagel and Stein [22] developed L^p-boundedness for a class of product singular integral operators. It was well known that any analysis of singular integrals on a product space $\widetilde{M} = M_1 \times \cdots \times M_n$ must be based on a formulation of standard singular integrals on each factor M_i. To carry this out, the important geometric objects used by Nagel and Stein are: (i) a class of equivalent control distances constructed on M_i, $1 \le i \le n$, via the vector fields $\{\mathbb{X}_1, \dots, \mathbb{X}_r\}$ where each $\mathbb{X}_i$, $1 \le i \le r$, depends on i; (ii) the volumes of balls satisfy the doubling property and the certain low bound estimates. More precisely, one variant of the control distance on M is defined as follows. For each $x, y \in M$, let $AC(x, y, \delta)$ denote the collection of absolutely continuous mapping $\varphi : [0, 1] \to M$ with $\varphi(0) = x$, $\varphi(1) = y$, and for almost every $t \in [0, 1]$, $\varphi'(t) = \sum_{j=1}^{k} a_j \mathbb{X}_j(\varphi(t))$ with $|a_j| \le \delta$. The control distance $\rho(x, y)$ from x to y is the infimum of the set of $\delta > 0$ such that $AC(x, y, \delta) \ne \varnothing$. See [22] and [24] for more details. It was shown in [22] that there is a pseudo-metric $d \approx \rho$ such that $d(x, y)$ is C^∞ on $M \times M \backslash \{\text{diagonal}\}$, and for $x \ne y$

$$|\partial_X^K \partial_Y^L d(x, y)| \lesssim d(x, y)^{1-K-L}.$$

Here ∂_X^K are products of K vector fields $\{X_1, \cdots X_r\}$ acting as derivatives on the x variable, and ∂_Y^L are corresponding K vector fields acting on the y variable.

The volume measure on M is defined in [22] as follows. One takes Lebesgue measure and denotes the measure of a set E by $|E|$. The ball is defined by $B(x, \delta) = \{y \in M, d(x, y) < \delta\}$ and the volume function is defined by $V(x, y) = |B(x, d(x, y))|$. Nagel and Stein proved that the volumes of the balls $B(x, \delta)$ satisfies the doubling property (see [22] for the details)

$$|B(x, 2\delta)| \le C|B(x, \delta)| \qquad \text{for all } \delta > 0 \text{ and some constant } C \tag{1.1}$$

and, moreover, it also satisfies the low bound condition, namely for $s \ge 1$, $|B(x, s\delta)| \ge s^4|B(x, \delta)|$ and for $s \le 1$, $|B(x, s\delta)| \approx s^4|B(x, \delta)|$.

We point out that the doubling condition (1.1) implies that there exist positive constants $C > 0$ and $\omega > 0$ (ω is said to be the upper dimension of M) such that for all $x \in M$ and $\lambda \ge 1$,

$$|B(x, \lambda r)| \lesssim C\lambda^\omega |B(x, r)|.$$

And the low bound condition implies the reverse doubling condition, that is, there are constants $\kappa \in (0, \omega]$ and $c \in (0, 1]$ such that

$$c\lambda^\kappa |(B(x, r))| \le |B(x, \lambda r)|$$

for all $x \in M$, $0 < r < \sup_{x,y \in M} d(x, y)/2$ and $1 \le \lambda < \sup_{x,y \in M} d(x, y)/2r$.

As it was pointed out in [22] there are two paths to formulate standard singular integrals on each factor M_i, $1 \le i \le n$. One is to generalize the class of operators on each factor M_i, $1 \le i \le n$, to the extended class of the $T1$ theorem of David and Journé [6] and then pass from this to a corresponding product theory. This was carried out in [17] for the setting where each factor is a Euclidean space. However, because of the inherent complications, Nagel and Stein [22] considered the class of singular integrals of NIS type; that is, non-isotropic smoothing operators of order 0. These operators may be viewed as Calderón-Zygmund operators whose kernels are C^∞ away from the diagonal and its cancellation conditions are given by their action on smooth bump functions. More precisely, on each factor M, these operators are defined by the following properties:

(I-1) If $\varphi, \psi \in C_0^\infty(M)$ have disjoint supports, then

$$\langle T\varphi, \psi\rangle = \int_{M\times M} K(x,y)\varphi(y)\psi(x)dydx.$$

(I-2) If φ is a normalized bump function associated to a ball of radius r, then $|\partial_X^a T\varphi| \lesssim r^{-a}$ for each integer $a \geq 0$.

(I-3) If $x \neq y$, then for every integer $a \geq 0$, $|\partial_{X,Y}^a K(x,y)| \lesssim d(x,y)^{-a}V(x,y)^{-1}$.

(I-4) Properties (I-1) through (I-3) also hold with x and y interchanged. That is, these properties also hold for the adjoint operator T^t defined by $\langle T^t\varphi, \psi\rangle = \langle T\psi, \varphi\rangle$.

To pass the above one factor case to the product theory, Nagel and Stein first consider two factors case $\widetilde{M} = M_1 \times M_2$. The product operator T on $\widetilde{M}$ is then defined from $C_0^\infty(\widetilde{M})$ to $C^\infty(\widetilde{M})$. The distribution $K(x_1, y_1, x_2, y_2)$, the Schwartz kernel of T, is a C^∞ function away from the "cross" $= \{(x,y) = (x_1, x_2, y_1, y_2) : x_1 = y_1$ and $x_2 = y_2\}$ and satisfies the following additional properties:

(II-1) $\big\langle T(\varphi_1\otimes\varphi_2), \psi_1\otimes\psi_2\big\rangle = \int K(x_1,y_1,x_2,y_2)\varphi_1(y_1)\varphi_2(y_2)\psi_1(x_1)\psi_2(x_2)dydx$ whenever $\varphi_i, \psi_i \in C_0^\infty(M_i)$ and have disjoint supports for $i = 1, 2$.

(II-2) For each bump function φ_2 on M_2 and each $x_2 \in M_2$, there exists a singular integral operator T^{φ_2,x_2} (of one parameter) on M_1, so that

$$\big\langle T(\varphi_1 \otimes \varphi_2), \psi_1 \otimes \psi_2\big\rangle = \int_{M_2} \big\langle T^{\varphi_2,x_2}\varphi_1, \psi_1\big\rangle \psi_2(x_2)dx_2.$$

Moreover, $x_2 \mapsto T^{\varphi_2,x_2}$ is smooth and uniform in the sense that T^{φ_2,x_2}, as well as $\rho_2^L \partial_{X_2}^L(T^{\varphi_2,x_2})$ for each $L \geq 0$, satisfy the conditions (I-1) to (I-4) uniformly.

(II-3) If φ_i is a bump function on a ball $B^i(r_i)$ in M_i, then

$$\big|\partial_{X_1}^{a_1}\partial_{X_2}^{a_2} T(\varphi_1 \otimes \varphi_2)\big| \lesssim r_1^{-a_1} r_2^{-a_2}.$$

In (II-2) and (II-3), both inequalities are taken in the sense of (I-2) whenever φ_2 is a bump function for $B^2(r_2)$ in M_2.

(II-4) $\big|\partial_{X_1,Y_1}^{a_1}\partial_{X_2,Y_2}^{a_2} K(x_1, y_1; x_2, y_2)\big| \lesssim \dfrac{d_1(x_1,y_1)^{-a_1} d_2(x_2,y_2)^{-a_2}}{V_1(x_1,y_1)V_2(x_2,y_2)}$.

(II-5) The same conditions hold when the index 1 and 2 are interchanged; that is, whenever the roles of M_1 and M_2 are interchanged.

(II-6) The same properties are assumed to hold for the 3 "transposes" of T, *i.e.*, those operators which arise by interchanging x_1 and y_1, or interchanging x_2 and y_2, or doing both interchanges.

The key to prove the L^2 and L^p, $1 < p < \infty$, boundedness of these operators in [22] is the existence of the Littlewood-Paley theory on $\widetilde{M}$, which is itself a consequence of the corresponding theory on each factor M_i, $1 \leq i \leq 2$. The square function used in [22] is constructed in terms of the heat equation. See [22] for the details.

Recently, in [14], inspired by the works of Nagel and Stein [21–23], the authors developed a satisfactory theory of multiparameter Hardy spaces in the framework of the product spaces of homogeneous type. Such a quasi-metric measure space of homogeneous type includes the model case of Carnot-Carathédory spaces intrinsic to a family of vector fields satisfying Hörmander's condition of finite rank. To be more precise, in [14] they consider (M, d, μ) to be a space of homogeneous type in the sense of Coifman and Weiis, that is, d is a quasi-metric satisfying (i) $d(x, y) = 0$ iff $x = y$; (ii) $d(x, y) = d(y, x)$; (iii) $d(x, z) \le A[d(x, y) + d(y, z)]$ for some $A \ge 1$. Moreover, $d(x, y)$ has the following regularity property

$$|d(x, y) - d(x', y)| \le C_0 d(x, x')^{\vartheta} [d(x, y) + d(x', y)]^{1-\vartheta} \tag{1.2}$$

for some regularity exponent $\vartheta : 0 < \vartheta \le 1$ and all $x,\ x',\ y \in M$. And μ is a nonnegative measure satisfying the following doubling and reverse doubling properties:

$$\mu(B(x, \lambda r)) \le C\lambda^{\omega} \mu(B(x, r)) \tag{1.3}$$

and

$$c\lambda^{\kappa} \mu((B(x, r))) \le \mu((B(x, \lambda r)) \tag{1.4}$$

where $B(x, r) = \{y : d(x, y) < r\}$ is the quasi-metric ball centered at x with radius r and $x \in M, 0 < r < \sup_{x,y\in M} d(x, y)/2$ and $1 \le \lambda < \sup_{x,y\in M} d(x, y)/2r$.

In [14] they introduced multiparametr Hardy spaces $H^p(\widetilde{M})$, provided the dual spaces of $H^p(\widetilde{M})$ in terms of multiparameter Carleson measure spaces $CMO^p(\widetilde{M})$, in particular, $BMO(\widetilde{M}) = CMO^1(\widetilde{M})$ and proved the endpoint estimates for those operators considered by Nagel and Stein in [22] on $H^p(\widetilde{M})$. See [14] for more details.

The main purpose of this paper is that under the same geometrical conditions as used in [14], we introduce a class of product operators, which generalizes Journé's class on the product Euclidean space in [17] and covers those studied by Nagel and Stein in [22]. Our goals are

(1) providing the criterion of the $L^2(\widetilde{M})$ boundedness for these general operators;
(2) proving the sharp endpoint estimates for these general operators on the multiparameter Hardy spaces $H^p(\widetilde{M})$ and the generalized Carleson measure spaces $CMO^p(\widetilde{M})$.

We now set our work in context. We begin with recalling $H^p(\widetilde{M})$ and $CMO^p(\widetilde{M})$ introduced in [14]. The crucial tool for developing $H^p(\widetilde{M})$ and $CMO^p(\widetilde{M})$ is the existence of a suitable approximation to the identity on one factor M. The construction of such an approximation to the identity is due to Coifman (see [7]). More precisely, take a smooth function h defined on $[0, \infty)$, equal to 1 on $[1, 2]$, and equal to 0 on $[0, 1/2]$ and on $[4, \infty)$. Let T_k be the operator with kernel $2^k h(2^k d(x, y))$ and M_k and W_k be the operators of multiplication by $1/T_k(1)$ and $\{T_k[1/T_k(1)]\}^{-1}$, respectively. Set $S_k := M_k T_k W_k T_k M_k$. The property (1.2) on the quasi-metric $d(x, y)$ and the conditions (1.3) and (1.4) on the measure μ imply that $S_k(x, y)$, the kernel of S_k, satisfy the following conditions: for some constants $C > 0$,

(i) $S_k(x, y) = 0$ for $d(x, y) \geq C2^{-k}$, and $|S_k(x, y)| \leq C \frac{1}{V_{2^{-k}}(x) + V_{2^{-k}}(y)}$,

(ii) $|S_k(x, y) - S_k(x', y)| \leq C \frac{2^{k\vartheta} d(x, x')^{\vartheta}}{V_{2^{-k}}(x) + V_{2^{-k}}(y)}$,

(iii) the above property (ii) also holds with x and y interchanged,

(iv) $\big|[S_k(x, y) - S_k(x, y')] - [S_k(x', y) - S_k(x', y')]\big| \leq C \frac{2^{2k\vartheta} d(x, x')^{\vartheta} d(y, y')^{\vartheta}}{V_{2^{-k}}(x) + V_{2^{-k}}(y)}$,

(v) $\int_X S_k(x, y)\, d\mu(y) = \int_X S_k(x, y)\, d\mu(x) = 1$,

where ϑ is same as in (1.2) and $V_r(x) := \mu(B(x, r))$.

The above sequence $\{S_k\}_{k\in\mathbb{Z}}$ of operators is said to be an approximation to the identity. To define the Littlewood-Paley-Stein square function, we also need to recall the spaces of test functions and distributions on M.

Definition 1.1 ([15]). Let $0 < \gamma, \beta \leq \vartheta$ where ϑ is the regularity exponent on M given in (1.2) and $r > 0$. A function f defined on M is said to be a test function of type (x_0, r, β, γ) centered at $x_0 \in M$ if f satisfies the following conditions

(i) $|f(x)| \leq C \frac{1}{V_r(x_0) + V(x, x_0)} \Big(\frac{r}{r + d(x, x_0)}\Big)^{\gamma}$;

(ii) $|f(x) - f(y)| \leq C \Big(\frac{d(x, y)}{r + d(x, x_0)}\Big)^{\beta} \frac{1}{V_r(x_0) + V(x, x_0)} \Big(\frac{r}{r + d(x, x_0)}\Big)^{\gamma}$ for all $x, y \in M$ with $d(x, y) \leq \frac{1}{2A}(r + d(x, x_0))$.

If f is a test function of type (x_0, r, β, γ), we write $f \in G(x_0, r, \beta, \gamma)$ and the norm of $f \in G(x_0, r, \beta, \gamma)$ is defined by

$$\|f\|_{G(x_0, r, \beta, \gamma)} = \inf\{C > 0 : \text{(i) and (ii) hold}\}.$$

Now fix $x_0 \in M$ we denote $G(\beta, \gamma) = G(x_0, 1, \beta, \gamma)$ and by $G_0(\beta, \gamma)$ the collection of all test functions in $G(\beta, \gamma)$ with $\int_M f(x) d\mu(x) = 0$. It is easy to check that $G(x_1, r, \beta, \gamma) = G(\beta, \gamma)$ with equivalent norms for all $x_1 \in M$ and $r > 0$. Furthermore, it is also easy to see that $G(\beta, \gamma)$ is a Banach space with respect to the norm in $G(\beta, \gamma)$.

Let $\overset{\circ}{G}_{\vartheta}(\beta, \gamma)$ be the completion of the space $G_0(\vartheta, \vartheta)$ in the norm of $G(\beta, \gamma)$ when $0 < \beta, \gamma < \vartheta$. If $f \in \overset{\circ}{G}_{\vartheta}(\beta, \gamma)$, we then define $\|f\|_{\overset{\circ}{G}_{\vartheta}(\beta,\gamma)} = \|f\|_{G(\beta,\gamma)}$. $\big(\overset{\circ}{G}_{\vartheta}(\beta, \gamma)\big)'$, the distribution space, is defined by the set of all linear functionals L from $\overset{\circ}{G}_{\vartheta}(\beta, \gamma)$ to $\mathbb{C}$ with the property that there exists $C \geq 0$ such that for all $f \in \overset{\circ}{G}_{\vartheta}(\beta, \gamma)$,

$$|L(f)| \leq C \|f\|_{\overset{\circ}{G}_{\vartheta}(\beta,\gamma)}.$$

Now we return to the product setting and recall the space of test functions and distributions on $\widetilde{M} = M_1 \times M_2$.

Definition 1.2 ([14]). Let $(x_0, y_0) \in \widetilde{M}$, $0 < \gamma_1, \gamma_2, \beta_1, \beta_2 \le \vartheta$ and $r_1, r_2 > 0$. A function $f(x, y)$ defined on $\widetilde{M}$ is said to be a test function of type $(x_0, y_0; r_1, r_2; \beta_1, \beta_2; \gamma_1, \gamma_2)$ if for any fixed $y, y' \in M_2$, $f(x, y)$, as a function of the variable of x, is a test function in $G^1(x_0, r_1, \beta_1, \gamma_1)$ on M_1. Similarly, for any fixed $x, x' \in M_1$, $f(x, y)$, as a function of the variable of y, is a test function in $G^2(y_0, r_2, \beta_2, \gamma_2)$ on M_2. Moreover, the following conditions are satisfied:

(i) $\|f(\cdot, y)\|_{G^1(x_0,r_1,\beta_1,\gamma_1)} \le C \dfrac{1}{V_{r_2}(y_0) + V(y_0, y)} \Big(\dfrac{r_2}{r_2 + d(y, y_0)}\Big)^{\gamma_2}$

(ii) $\|f(\cdot, y) - f(\cdot, y')\|_{G^1(x_0,r_1,\beta_1,\gamma_1)}$
$$\le C\Big(\frac{d(y, y')}{r_2 + d(y, y_0)}\Big)^{\beta_2} \frac{1}{V_{r_2}(y_0) + V(y_0, y)} \Big(\frac{r_2}{r_2 + d(y, y_0)}\Big)^{\gamma_2}$$
for all $y, y' \in M_2$ with $d(y, y') \le (r_2 + d(y, y_0))/2A_2$;

(iii) Both properties (i) and (ii) also hold with x, y and G^1, G^2 interchanged.

If f is a test function of type $(x_0, y_0; r_1, r_2; \beta_1, \beta_2; \gamma_1, \gamma_2)$, we write $f \in G(x_0, y_0; r_1, r_2; \beta_1, \beta_2; \gamma_1, \gamma_2)$ and the norm of f is defined by

$$\|f\|_{G(x_0,y_0;r_1,r_2;\beta_1,\beta_2;\gamma_1,\gamma_2)} = \inf\{C : \text{(i), (ii) and (iii) hold}\}.$$

Similarly, we denote by $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ the class of $G(x_0, y_0; 1, 1; \beta_1, \beta_2; \gamma_1, \gamma_2)$ for any fixed $(x_0, y_0) \in \widetilde{M}$. Set that $f(x, y) \in G_0(\beta_1, \beta_2; \gamma_1, \gamma_2)$ if $\int_{M_1} f(x, y)d\mu^1(x) = \int_{M_2} f(x, y)d\mu^2(y) = 0$. We can check that $G(x_0, y_0; r_1, r_2; \beta_1, \beta_2; \gamma_1, \gamma_2) = G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ with equivalent norms for all $(x_0, y_0) \in \widetilde{M}$ and $r_1, r_2 > 0$. Furthermore, it is easy to see that $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ is a Banach space with respect to the norm in $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$.

Let $\mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ be the completion of the space $G_0(\vartheta_1, \vartheta_2; \vartheta_1, \vartheta_2)$ in $G(\beta_1, \beta_2; \gamma_1, \gamma_2)$ with $0 < \beta_i, \gamma_i < \vartheta_i$, where ϑ_i is the regularity exponent on M_i, $i = 1, 2$. If $f \in \mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$, we then define $\|f\|_{\mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1,\beta_2;\gamma_1,\gamma_2)} = \|f\|_{G(\beta_1,\beta_2;\gamma_1,\gamma_2)}$.

We define the distribution space $\big(\mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)\big)'$ by all linear functionals L from $\mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ to $\mathbb{C}$ with the property that there exists $C \ge 0$ such that for all $f \in \mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$,

$$|L(f)| \le C\|f\|_{\mathring{G}_{\vartheta_1,\vartheta_2}(\beta_1,\beta_2;\gamma_1,\gamma_2)}.$$

We now recall the Littlewood-Paley-Stein square function, the Hardy space and the generalized Carleson measure space on $\widetilde{M} = M_1 \times M_2$.

Definition 1.3 ([14]). Let $\{S^i_{k_i}\}_{k_i\in\mathbb{Z}}$ be approximations to the identity on M_i and $D^i_{k_i} = S^i_{k_i} - S^i_{k_i-1}$, $i = 1, 2$. For $f \in \big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)\big)'$ with $0 < \beta_i, \gamma_i < \vartheta_i$, $i = 1, 2$, $S(f)$, the Littlewood-Paley-Stein square function of f, is defined by

$$S(f)(x, y)=\left\{\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{\tau_1}\sum_{\tau_2}|D^1_{k_1}D^2_{k_2}(f)(x, y)|^2\chi_{I^{k_1+N_1}_{1,\tau_1}}(x)\chi_{I^{k_2+N_2}_{2,\tau_2}}(y)\right\}^{1/2},$$

where $I^{k_i+N_i}_{i,\tau_i}$, $i = 1, 2$, are "dyadic cubes" in M_i in the sense of Christ [4] (see Theorem 2.1 below and also [27]).

Definition 1.4 ([14]). Let $\max\big(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\big) < p \le 1$, $0 < \beta_i, \gamma_i < \vartheta_i$ for $i = 1, 2$, and N_1, N_2 are fixed large integers. Let $\{S^i_{k_i}\}_{k_i\in\mathbb{Z}}$ be an approximation to the identity on M_i and for $k_i \in \mathbb{Z}$, set $D^i_{k_i} = S^i_{k_i} - S^i_{k_i-1}$, $i = 1, 2$. The Hardy space $H^p(\widetilde{M})$ is defined by the set of all $f \in \big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)\big)'$ such that

$$H^p(\widetilde{M}) =\left\{f\in\big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)\big)' : 0<\beta_i, \gamma_i<\vartheta_i, i=1, 2, S(f)\in L^p(\widetilde{M})\right\}$$

and if $f \in H^p(\widetilde{M})$, the norm of f is defined by $\|f\|_{H^p(\widetilde{M})} = \|S(f)\|_p$.

Definition 1.5 ([14]). Let $\max\big(\frac{2\omega_1}{2\omega_1+\vartheta_1}, \frac{2\omega_2}{2\omega_2+\vartheta_2}\big) < p \le 1$, $0 < \beta_i, \gamma_i < \vartheta_i$ for $i = 1, 2$, and N_1, N_2 are fixed large integers. Let $\{S^i_{k_i}\}_{k_i\in\mathbb{Z}}$ be an approximation to the identity on M_i and for $k_i \in \mathbb{Z}$, set $D^i_{k_i} = S^i_{k_i} - S^i_{k_i-1}$, $i = 1, 2$. The generalized Carleson measure space $CMO^p(\widetilde{M})$ is defined by the set of all $f \in \big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)\big)'$ such that

$$\begin{aligned}&\|f\|_{CMO^p(\widetilde{M})}\\ &=\sup_{\Omega}\left\{\frac{1}{|\Omega|^{\frac{2}{p}-1}}\int_{\Omega}\sum_{k_1,k_2}\sum_{I_1\times I_2\subseteq\Omega}\big|D^1_{k_1}D^2_{k_2}(f)(x, y)\big|^2\chi_{I_1}(x)\chi_{I_2}(y)d\mu^1(x)d\mu^2(y)\right\}^{\frac{1}{2}}\\ &<\infty,\end{aligned}$$

where Ω ranges over all open sets in $\widetilde{M}$ with finite measures and for each k_1 and k_2, I_1, I_2 range over all the dyadic cubes in M_1 and M_2 with length $\ell(I_1) = 2^{-k_1-N_1}$ and $\ell(I_2) = 2^{-k_2-N_2}$, respectively.

To see why these definitions are well posed, and in particular, why fixed large integers N_1 and N_2 are needed, the crucial tool is the following Carlderón reproducing formula on $\widetilde{M}$.

Let $\{S^i_{k_i}\}_{k_i\in\mathbb{Z}}$ be approximations to the identity on M_i and $D^i_{k_i} = S^i_{k_i} - S^i_{k_i-1}$, $i = 1, 2$. Then the Calderón reproducing formula is given by

$$\begin{aligned} f(x, y) = &\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{\tau_1}\sum_{\tau_2}\mu_1(I^{k_1+N_1}_{1,\tau_1})\mu_2(I^{k_2+N_2}_{2,\tau_2}) \\ &\times \widetilde{D^1}_{k_1}(x, y^{k_1+N_1}_{\tau_1})\widetilde{D^2}_{k_2}(y, y^{k_2+N_2}_{\tau_2})D^1_{k_1}D^2_{k_2}(f)(y^{k_1+N_1}_{\tau_1}, y^{k_2+N_2}_{\tau_2}) \\ = &\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{I_1}\sum_{I_2}\mu_1(I_1)\mu_2(I_2) \\ &\times D^1_{k_1}(x_1, x_{I_1})D^2_{k_2}(x_2, x_{I_2})\widetilde{\widetilde{D^1}}_{k_1}\widetilde{\widetilde{D^2}}_{k_2}(f)(x_{I_1}, x_{I_2}), \end{aligned} \tag{1.5}$$

where the series converges in both the norm of $\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta'_1, \beta'_2, \gamma'_1, \gamma'_2)$ with $0 < \beta'_i < \beta_i < \vartheta_i, \gamma'_i, \gamma_i < \vartheta_i, i = 1, 2$, and the norm of $L^p(M_1 \times M_2)$, $1 < p < \infty$. See [14] for the existence of operators $\widetilde{D^1}_{k_1}, \widetilde{D^2}_{k_2}, \widetilde{\widetilde{D^1}}_{k_1}, \widetilde{\widetilde{D^2}}_{k_2}$ and the choice of fixed N_1 and N_2.

It was well known that this kind of identities is a powerful tool in classical harmonic analysis. See [1] and [20] for the classical case and [5] for spaces of homogeneous type. Applying the above Carlderón reproducing formula, in [14] it was proved that for $\max\left(\frac{2\omega_1}{2\omega_1+\vartheta_1}, \frac{2\omega_2}{2\omega_2+\vartheta_2}\right) < p \le 1$,

$$\left(H^p(\widetilde{M})\right)' = CMO^p(\widetilde{M}).$$

In particular,

$$\left(H^1(\widetilde{M})\right)' = CMO^1(\widetilde{M}) = BMO(\widetilde{M}).$$

Now we are ready to introduce a class of general singular integral operators on $\widetilde{M} = M_1 \times M_2$ and state our main results.

Let $C^\eta_0(M_1)$ denote the space of continuous functions f with compact support such that

$$\|f\|_{\eta(M_1)} := \sup_{x,y\in M_1, x\neq y}\frac{|f(x) - f(y)|}{d_1(x, y)^\eta} < \infty$$

and $C^\eta_0(M_2)$ is defined similarly. And let $C^\eta_0(\widetilde{M}), \eta > 0$, denote the space of continuous functions f with compact support such that

$$\|f\|_\eta := \sup_{x_1\neq y_1, x_2\neq y_2}\frac{|f(x_1, x_2) - f(y_1, x_2) - f(x_1, y_2) + f(y_1, y_2)|}{d_1(x_1, y_1)^\eta d_2(x_2, y_2)^\eta} < \infty.$$

We first consider one factor case. A continuous function $K_1(x_1, y_1)$ defined on $M_1\backslash\{(x_1, y_1) : x_1 = y_1\}$ is called a *Calderón-Zygmund kernel* if there exist constant

$C > 0$ and a regularity exponent $\varepsilon_1 \in (0, 1]$ such that

(a) $|K_1(x_1, y_1)| \le C V(x_1, y_1)^{-1}$;

(b) $|K_1(x_1, y_1) - K_1(x_1, y_1')| \le C\Big(\dfrac{d_1(y_1, y_1')}{d_1(x_1, y_1)}\Big)^{\varepsilon_1} V(x_1, y_1)^{-1}$
if $d_1(y_1, y_1') \le \dfrac{d_1(x_1, y_1)}{2A_1}$;

(c) $|K_1(x_1, y_1) - K_1(x_1', y_1)| \le C\Big(\dfrac{d_1(x_1, x_1')}{d_1(x_1, y_1)}\Big)^{\varepsilon_1} V(x_1, y_1)^{-1}$
if $d_1(x_1, x_1') \le \dfrac{d_1(x_1, y_1)}{2A_1}$.

The smallest such constant C is denoted by $|K_1|_{CZ}$. We say that an operator T_1 is a *singular integral operator* associated with a Calderón-Zygmund kernel K_1 if the operator T_1 is a continuous linear operator from $C_0^\eta(M_1)$ into its dual such that

$$\langle T_1 f, g\rangle = \iint g(x_1)K_1(x_1, y_1) f(y_1)d\mu^1(y_1)d\mu^1(x_1)$$

for all functions $f, g \in C_0^\eta(M_1)$ with disjoint supports. T_1 is said to be a *Calderón-Zygmund operator* if it extends to be a bounded operator on $L^2(M_1)$. If T_1 is a Calderón-Zygmund operator associated with a kernel K_1, its operator norm is defined by $\|T_1\|_{CZ} = \|T_1\|_{L^2\to L^2} + |K_1|_{CZ}$.

Now we introduce a class of the *product singular integral operators* on $\widetilde{M}$. Let $T : C_0^\eta(\widetilde{M}) \to [C_0^\infty(\widetilde{M})]'$ be a linear operator. T is said to be a singular integral operator if there exists a pair (K_1, K_2) of Calderón-Zygmund valued operators on M_2 and M_1, respectively, such that

$$\langle g\otimes k, Tf\otimes h\rangle = \iint g(x_1)\langle k, K_1(x_1, y_1)h\rangle f(y_1)d\mu^1(x_1)d\mu^1(y_1)$$

for all $f, g \in C_0^\eta(M_1)$ and $h, k \in C_0^\eta(M_2)$, with $\operatorname{supp} f \cap \operatorname{supp} g = \varnothing$ and

$$\langle k\otimes g, Th\otimes f\rangle = \iint g(x_2)\langle k, K_2(x_2, y_2)h\rangle f(y_2)d\mu^2(x_2)d\mu^2(y_2)$$

for all $f, g \in C_0^\eta(M_2)$ and $h, k \in C_0^\eta(M_1)$, with $\operatorname{supp} f \cap \operatorname{supp} g = \varnothing$. Moreover, $\|K_i(x_i, y_i)\|_{CZ}$, $i = 1, 2$, as functions of $x_i, y_i \in M_i$, satisfy the following conditions:

(i) $\|K_i(x_i, y_i)\|_{CZ} \le C V(x_i, y_i)^{-1}$;

(ii) $\|K_i(x_i, y_i) - K_i(x_i, y_i')\|_{CZ} \le C\Big(\dfrac{d_i(y_i, y_i')}{d_i(x_i, y_i)}\Big)^{\varepsilon_i} V(x_i, y_i)^{-1}$
if $d_i(y_i, y_i') \le \dfrac{d_i(x_i, y_i)}{2A_i}$;

(iii) $\|K_i(x_i, y_i) - K_i(y_i', y_i)\|_{CZ} \le C\left(\frac{d_i(x_i, y_i')}{d_i(x_i, y_i)}\right)^{\varepsilon_i} V(x_i, y_i)^{-1}$
if $d_i(x_i, y_i') \le \frac{d_i(x_i, y_i)}{2A_i}$.

We remark, as mentioned, that the above class of the product singular integral operators includes those introduced by Journé in [17] and studied by Nagel and Stein in [22].

Suppose that T is such a product singular integral operator on $\widetilde{M}$. T is said to be a *product Calderón-Zygmund operator* on $\widetilde{M}$ if T extends to be a bounded operator on L^2.

We now describe how a product singular integral operator T acts on bounded $C^\eta(\widetilde{M})$ functions (denote by $C_b^\eta(\widetilde{M})$). Following Journé in [17], we first define the operator T_1 by the following

$$\langle g_1 \otimes g_2, Tf_1 \otimes f_2\rangle = \langle g_2, \langle g_1, T_1 f_1\rangle f_2\rangle$$

for $f_1, g_1 \in C_0^\eta(M_1)$ and $f_2, g_2 \in C_0^\eta(M_2)$. Note that when $g_1 \in C_{00}^\eta(M_1)$ and $f_1 \in C_b^\eta(M_1)$, the inner product $\langle g_1, T_1 f_1\rangle$ is well defined. Moreover, $\langle g_1, T_1 f_1\rangle$ is a singular integral operator on M_2 with a Calderón-Zygmund kernel $\langle g_1, T_1 f_1\rangle(x_2, y_2) = \langle g_1, K_2(x_2, y_2) f_1\rangle$. Therefore, for $g_2 \in C_{00}^\eta(M_2)$ and $f_2 \in C_b^\eta(M_2)$, $\langle g_2, \langle g_1, T_1 f_1\rangle f_2\rangle$ is well defined. One defines $\langle g_2, T_2 f_2\rangle$ similarly for $g_2 \in C_{00}^\eta(M_2)$ and $f_2 \in C_b^\eta(M_2)$. Using these definitions, we can give a meaning of the notation $T1 = 0$. More precisely, $T1 = 0$ means $\langle g_1 \otimes g_2, T1\rangle = 0$ for all $g_1 \in C_{00}^\eta(M_1)$ and $g_2 \in C_{00}^\eta(M_2)$, that is,

$$\iint g(x_1)g(x_2)K(x_1, x_2, y_1, y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1)d\mu^2(y_2) = 0.$$

Similarly, $T_1(1) = 0$ is equivalent to $\langle g_1, \langle g_2, T_2 f_2\rangle 1\rangle = 0$ for all $g_1 \in C_{00}^\eta(M_1)$ and $f_2, g_2 \in C_0^\eta(M_2)$, that is, for $g_1 \in C_{00}^\eta(M_1)$, $g_2 \in C_{00}^\eta(M_2)$ and almost everywhere $y_2 \in M_2$,

$$\iint g(x_1)g(x_2)K(x_1, x_2, y_1, y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1) = 0.$$

While $T_1{}^*(1) = 0$ means $\langle g_2, T_2 f_2\rangle^* 1 = 0$ in the same conditions. Interchanging the role of indices one obtains the meaning of $T_2(1) = 0$ and $T_2{}^*(1) = 0$.

We also need to introduce the definition of weak boundedness property (denote by WBP). We begin with the one factor case. Let T_1 be a singular integral operator on M_1 and let $A_{M_1}(\delta, x_1^0, r_1)$, $\delta \in (0, \vartheta_1]$, $x_1^0 \in M_1$ and $r_1 > 0$, be a set of all $f \in C_0^\delta(M_1)$ supported in $B(x_1^0, r_1)$ satisfying $\|f\|_\infty \le 1$ and $\|f\|_\delta \le r_1^{-\delta}$. We say that T_1 has the weak boundedness property if there exist $0 < \delta \le \vartheta_1$ and a constant $C > 0$ such that for all $x_1^0 \in M_1$, $r_1 > 0$, and all $\phi, \psi \in A_{M_1}(\delta, x_1^0, r_1)$,

$$|\langle T_1\phi, \psi\rangle| \le C V_{r_1}(x_1^0).$$

Similarly we can define the set $A_{M_2}(\delta, x_2^0, r_2)$, $\delta \in (0, \vartheta_2]$, $x_2^0 \in M_2$ and the weak boundedness property for a singular integral operator on M_2.

In the following, we define the weak boundedness property in the product setting.

Definition 1.6. Let T be a product singular integral operator on $\widetilde{M}$. T has the WBP if

$$\|\langle T_1\phi^1, \psi^1\rangle\|_{CZ} \lesssim V_{r_1}(x_1^0) \qquad \text{for all } \phi^1, \psi^1 \in A_{M_1}(\delta, x_1^0, r_1), \tag{1.6}$$

$$\|\langle T_2\phi^2, \psi^2\rangle\|_{CZ} \lesssim V_{r_2}(x_2^0) \qquad \text{for all } \phi^2, \psi^2 \in A_{M_2}(\delta, x_2^0, r_2). \tag{1.7}$$

It is easy to see that if T satisfies (1.6) and (1.7), then

$$|\langle T\phi^1 \otimes \phi^2, \psi^1 \otimes \psi^2\rangle| \lesssim V_{r_1}(x_1^0)V_{r_2}(x_2^0)$$

for all $\phi^1, \psi^1 \in A_{M_1}(\delta, x_1^0, r_1)$ and $\phi^2, \psi^2 \in A_{M_2}(\delta, x_2^0, r_2)$.

We point out that if T is a product Calderón-Zygmund operator on $\widetilde{M}$, then T has the weak boundedness property. We denote $\widetilde{T}$ by the partial adjoint operator of T with the kernel $K(y_1, x_2, x_1, y_2)$.

Main results of this paper are the following:

Theorem 1.7. *Let T be a product singular integral operator on $\widetilde{M}$. Then T and $\widetilde{T}$ are both bounded on $L^2(\widetilde{M})$ if and only if $T1$, T^*1 $\widetilde{T}1$, and $(\widetilde{T})^*1$ lie in $BMO(\widetilde{M})$ and T has the weak boundedness property.*

Theorem 1.8. *Let T be a product Calderón-Zygmund operator on $\widetilde{M}$ and $\max\left(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\right) < p \le 1$. Then T extends to a bounded operator from $H^p(\widetilde{M})$ to itself if and only if $(T^*)_1(1) = (T^*)_2(1) = 0$.*

Theorem 1.9. *Let T be a product Calderón-Zygmund operator on $\widetilde{M}$ and $\max\left(\frac{2\omega_1}{2\omega_1+\vartheta_1}, \frac{2\omega_2}{2\omega_2+\vartheta_2}\right) < p \le 1$. Then T extends to a bounded operator from $CMO^p(\widetilde{M})$ to itself, particularly from $BMO(\widetilde{M})$ to itself, if and only if $T_1(1) = T_2(1) = 0$.*

Before ending this section, we would like to describe our strategy of the proofs. To show the necessity of Theorem 1, we will employ an approach which is different from the one given by Journé [17]. Note that Journé obtained this implication by showing that the L^2 boundedness implies the $L^\infty \to BMO$ boundedness. For this purpose, he established a fundamental geometric covering lemma. As a consequence of this implication, together with an interpolation theorem and the duality argument, Journé proved that the L^2 boundedness implies the L^p, $1 < p < \infty$, boundedness. We will prove this implication by applying the Hardy space theory developed in [14]. More precisely, we first show that the L^2 boundedness implies the $H^1 \to L^1$ boundedness. We would like to point out that the $H^1 \to L^1$ boundedness was obtained in [14] under the cancellation conditions used in [22]. However, this is not available for the current situation. To show that the L^2 boundedness implies the $H^1 \to L^1$ boundedness without assuming any cancellation conditions, we will apply an atomic decomposition for $H^p(\widetilde{M})$. For this purpose, following Pipher's idea [25], we first establish Journé-type covering lemma and develop an atomic decomposition in our setting. Applying an atomic decomposition and a similar idea as in [9], we conclude that L^2 boundedness implies the $H^p \to L^p$ boundedness, particularly, $H^1 \to L^1$ boundedness. From this together with the duality between $H^1(\widetilde{M})$ and $BMO(\widetilde{M})$ we obtain the $L^\infty \to BMO$ boundedness and hence the desired necessary condition of Theorem 1 follows. Moreover, by an interpolation theorem proved in [14], we also conclude that the L^2 boundedness implies the L^p, $1 < p < \infty$, boundedness.

In [17] the proof of the sufficiency of the classical product $T1$ theorem was decomposed in three steps. In the first step, Journé claimed that if T satisfies $T_1(1) = T_1^*(1) = 0$ and has the weak boundedness property, then it can be viewed as a classical vector valued singular integral operator acting on $C_0^\infty(\mathbb{R}) \times H$ and the L^2-boundedness of such an operator follows from the classical case. The second step is the decomposition of an operator T having the weak boundedness properties and $T(1) = T^*(1) = \widetilde{T}(1) = \widetilde{T}^*(1) = 0$ as the sum of two operators S and $T - S$ with both the weak boundedness properties and $S_2(1) = S_2^*(1) = (T - S)_1(1) = (T - S)_1^*(1) = 0$. The L^2 boundedness of T is then a consequence of the first step. The last step is, as in the classical one parameter case, to construct the para-product operator W_b, for $b \in BMO(R \times R)$ so that $W_b 1 = b$, $W_b^* 1 = \widetilde{W}_b 1 = \widetilde{W}_b^* 1 = 0$. If one sets $S = T - W_{T1} - W_{T^*1}^* - \widetilde{W}_{\widetilde{T}1} - \widetilde{W}_{\widetilde{T}^*1}^*$, then $S(1) = S^*(1) = \widetilde{S}(1) = \widetilde{S}^*(1) = 0$. Moreover, all para-product operators W_b, W_b^*, $\widetilde{W}_b$ and $\widetilde{W}_b^*$ are in Journé's class and bounded on $L^2(\mathbb{R} \times \mathbb{R})$. See [17] for all details. We will develop a new approach to prove the sufficiency of Theorem 1. To describe ideas of this approach, we first outline a new proof for the classical $T1$ theorem on M_1. As in the classical case, we consider the following bilinear form

$$\langle g, Tf\rangle = \left\langle \sum_{j=-\infty}^{\infty} D_j \widetilde{D}_j(g), T \sum_{k=-\infty}^{\infty} D_k \widetilde{D}_k(f) \right\rangle = \sum_{j,k} \langle \widetilde{D}_j(g), D_j T D_k \widetilde{D}_k(f)\rangle.$$

The original proof of the classical $T1$ theorem includes two steps. In the first step, if T is a singular integral operator having the weak boundedness property and $T(1) = T^*(1) = 0$, then $D_j T D_k(x, y)$, the kernel of the operator $D_j T D_k$, satisfies the almost orthogonal estimate. From this together with the Littlewood-Paley estimate on L^2 implies $|\langle g, Tf\rangle| \le C\|f\|_2\|g\|_2$ and hence the L^2 boundedness of T follows. The second step is to reduce the general case to the first case in terms of the para-product operator. Observing that without assuming any cancellation condition on K the following almost orthogonal estimate for $j \le k$ still holds

$$\left| \iint [D_j(x, u) - D_j(x, y)] K(u, v) D_k(v, y) d\mu^1(u) d\mu^1(v) \right|$$
$$\le C 2^{(j-k)\epsilon} \frac{1}{V_{2^{-j}}(x) + V_{2^{-j}}(y) + V(x, y)} \frac{2^{-j\varepsilon}}{(2^{-j} + d(x, y))^{\varepsilon}},$$

and similarly for $k < j$. This leads to the following decomposition:

$$
\begin{aligned}
\langle g, Tf\rangle =&\sum_{j\le k}\int \widetilde{D}_j(g)(x)\iint [D_j(x,u)-D_j(x,y)]K(u,v)\\
&\quad\times D_k(v,y)d\mu^1(u)d\mu^1(v)\widetilde{D}_k(f)(y)d\mu^1(y)d\mu^1(x)\\
&+\sum_{k<j}\int \widetilde{D}_j(g)(x)\iint D_j(x,u)K(u,v)\\
&\quad\times [D_k(v,y)-D_k(x,y)]d\mu^1(u)d\mu^1(v)\widetilde{D}_k(f)(y)d\mu^1(y)d\mu^1(x)\\
&+\sum_{j\le k}\int \widetilde{D}_j(g)(x)\iint D_j(x,y)K(u,v)\\
&\quad\times D_k(v,y)d\mu^1(u)d\mu^1(v)\widetilde{D}_k(f)(y)d\mu^1(y)d\mu^1(x)\\
&+\sum_{k<j}\int \widetilde{D}_j(g)(x)\iint D_j(x,u)K(u,v)\\
&\quad\times D_k(x,y)d\mu^1(u)d\mu^1(v)\widetilde{D}_k(f)(y)d\mu^1(y)d\mu^1(x).
\end{aligned}
$$

The almost orthogonal estimates, as mentioned above, together with the Littlewood-Paley estimate on L^2 imply that the first two series are bounded by some constant $C\|f\|_2\|g\|_2$. To see that the last two series are also bounded by $C\|f\|_2\|g\|_2$, we only consider the third series and rewrite it as $\int\sum_k \widetilde{S}_k(g)(y)D_k(T^*1)(y)\cdot \widetilde{D}_k(f)(y)d\mu^1(y)$, where $\widetilde{S}_k=\sum_{j\le k}D_j\widetilde{D}_j$. The Carleson measure estimate together with Littlewood-Paley estimate yields

$$
\begin{aligned}
&\Big|\int\sum_k \widetilde{S}_k(g)(y)D_k(T^*1)(y)\widetilde{D}_k(f)(y)d\mu^1(y)\Big|\\
\le&\Big\{\int\sum_k|\widetilde{S}_k(g)(y)|^2|D_k(T^*1)(y)|^2d\mu^1(y)\Big\}^{\frac12}\Big\{\int\sum_k|\widetilde{D}_k(f)(y)|^2d\mu^1(y)\Big\}^{\frac12}\\
\le&C\|f\|_2\|g\|_2.
\end{aligned}
$$

This new approach can be carried out to the product case. Indeed, by the Calderón's reproducing formula in (1.5) on the product $\widetilde{M}$, we begin with the following bilinear form

$$
\begin{aligned}
\langle g, Tf\rangle =&\sum_{k_1'}\sum_{\tau_1'}\sum_{k_1}\sum_{\tau_1}\sum_{k_2'}\sum_{\tau_2'}\sum_{k_2}\sum_{\tau_2}\mu((I^1)')\mu(I^1)\mu((I^2)')\mu(I^2)\\
&\times \widetilde{D^1}_{k_1'}\widetilde{D^2}_{k_2'}(g)(x_{(I^1)'},x_{(I^2)'})\\
&\times \big\langle D^1_{k_1'}D^2_{k_2'}, TD^1_{k_1}D^2_{k_2}\big\rangle(x_{(I^1)'},x_{(I^2)'},x_{I^1},x_{I^2})\widetilde{D^1}_{k_1}\widetilde{D^2}_{k_2}(f)(x_{I^1},x_{I^2}).
\end{aligned}
$$

We would also like to point out that

$$\begin{aligned}\langle D^1_{k_1'} D^2_{k_2'}, T D^1_{k_1} D^2_{k_2}\rangle &= \langle D^1_{k_1'}, \langle D^2_{k_2'}, K_1(x_1, y_1) D^2_{k_2}\rangle D^1_{k_1}\rangle \\ &= \langle D^2_{k_2'}, \langle D^1_{k_1'}, K_2(x_2, y_2) D^1_{k_1}\rangle D^2_{k_2}\rangle,\end{aligned}$$

which will be crucial for this new approach.

Our strategy of the proof of the sufficiency of Theorem 1 uses a similar decomposition as one parameter case. However, some new mixed situations have to be taken into account. More precisely, except applying the almost orthogonal estimate and Carleson measure estimate on $M_1 \times M_2$, one also needs to consider two more mixed cases: the almost orthogonal estimate on one factor, say M_1, and Carleson measure estimate on M_2, and the Littlewood-Paley estimate on one factor, say M_1, and Carleson measure estimate on M_2. See more details in Subsection 2.2.

Note that in [17] Journé proved that if T is a convolution operator and bounded on L^2, then T admits a bounded extension from $BMO(\mathbb{R} \times \mathbb{R})$ to itself. He mentioned without proof that if T is a Calderón-Zygmund operator and $T_1(1) = T_2(1) = 0$, then TH_1, TH_2 and TH_1H_2 are Cadelrón-Zygmund operators, where H_1, H_2 and H_1H_2 are the Hilbert transforms and double Hilbert transform. From this together with the characterization of the product $BMO(\mathbb{R} \times \mathbb{R})$ in terms of the bi-Hilbert transform, the boundedness of T on $BMO(\mathbb{R} \times \mathbb{R})$ follows. In our setting, however, his method is not available. In this paper, the L^2 theory and the duality argument between $H^p(\widetilde{M})$ and $CMO^p(\widetilde{M})$ will play a crucial role in the proofs of Theorems 2 and 3. More precisely, it is known that $L^2(\widetilde{M}) \cap H^p(\widetilde{M})$ is dense in $H^p(\widetilde{M})$. Therefore, to show that Tf is bounded on $H^p(\widetilde{M})$, it suffices to consider $f \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$. This argument for space $CMO^p(\widetilde{M})$ is no long true. However, $L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$ is dense in the weak topology (H^p, CMO^p). Then applying this density argument together with the duality argument implies the sufficiency of Theorem 3. We will show the necessity of Theorem 3 first and the same conclusion for Theorem 2 follows from the density argument and the duality argument. This approach is new even for the classical case.

The paper is organized as follows. In Section 2, we we prove Theorem 1. In Subsection 2.2, we prove the necessity. Journé-type covering lemma and atomic decomposition are provided in Subsections 2.2.1 and 2.2.2. We prove that if T is bounded on L^2 then T extends to a bounded operator from H^p to L^p, from L^∞ to BMO, and from L^p to itself in Subsections 2.2.3, 2.2.4 and Subsection 2.2.5, respectively. The sufficiency of Theorem 1 is proved in the Subsection 2.3. The proofs of Theorem 2 and 3 will be given in Section 3. In the last section, we will point out that all results in this paper can be carried out to the case with arbitrarily many parameters. We, however, state these results only and omit the details of the proofs.

Throughout the paper, $A \approx B$ means that the ratio A/B is bounded and bounded away from zero by constants that do not depend on the relevant variables in A and B. $A \lesssim B$ means that the ratio A/B is bounded by a constant independent of the relevant variables.

2. Proof of Theorem 1

In this section we prove Theorem 1.

2.1. Necessity of Theorem 1

To show the necessity of Theorem 1, we will employ the Hardy space theory on $\widetilde{M}$ developed in [14]. As mentioned in Section 1, we first show that if T is a Calderón-Zygmund operator on $\widetilde{M}$ then T extends to a bounded operator from $H^p(\widetilde{M})$ to $L^p(\widetilde{M})$ for $p \le 1$ and is close to 1. This, particularly for $p = 1$, together with the duality (L^1, L^∞) and (H^1, BMO), implies that T is bounded from L^∞ to BMO. To achieve this goal, the main tool we need is an atomic decomposition for $H^p(\widetilde{M})$. To this end, as in the classical case, we shall first provide Journé-type covering lemma on $\widetilde{M}$, for which we turn to next subsection.

2.1.1. *Journé-type covering lemma*

We first need a result of Christ.

Theorem 2.1 ([4]). *Let (M, ρ, μ) be a space of homogeneous type, then, there exists a collection $\{I_\alpha^k \subset M : k \in \mathbb{Z}, \alpha \in I^k\}$ of open subsets, where I^k is some index set, and $C_1, C_2 > 0$, such that*

(i) $\mu(M \setminus \bigcup_\alpha I_\alpha^k) = 0$ *for each fixed k and $I_\alpha^k \bigcap I_\beta^k = \varnothing$ if $\alpha \ne \beta$;*

(ii) *for any α, β, k, l with $l \ge k$, either $I_\beta^l \subset I_\alpha^k$ or $I_\beta^l \bigcap I_\alpha^k = \varnothing$;*

(iii) *for each (k, α) and each $l < k$ there is a unique β such that $I_\alpha^k \subset I_\beta^l$;*

(iv) $\operatorname{diam}(I_\alpha^k) \le C_1 2^{-k}$*;*

(v) *each I_α^k contains some ball $B(z_\alpha^k, C_2 2^{-k})$, where $z_\alpha^k \in M$.*

Note that Carnot-Carathéodory spaces are spaces of homogeneous type. Therefore, we can think of I_α^k as being a dyadic cube with diameter rough 2^{-k} centered at z_α^k. As a result, we consider CI_α^k to be the cube with the same center as I_α^k and diameter $C\operatorname{diam}(I_\alpha^k)$. To simplify notations, we will call I dyadic cubes and denote the side length of I by $\ell(I)$.

Let $\{I_{\tau_i}^{k_i} \subset M_i : k_i \in \mathbb{Z}, \tau_i \in I^{k_i}\}$ be the same as in Theorem 2.1. We call $R = I_{\tau_1}^{k_1} \times I_{\tau_2}^{k_2}$ a dyadic rectangle in $\widetilde{M}$. Let $\Omega \subset \widetilde{M}$ be an open set of finite measure and $\mathcal{M}_i(\Omega)$ denote the family of dyadic rectangles $R \subset \Omega$ which are maximal in the ith "direction", $i = 1, 2$. Also we denote by $\mathcal{M}(\Omega)$ the set of all maximal dyadic rectangles contained in Ω. For the sake of simplicity, we denote by $R = I_1 \times I_2$

any dyadic rectangles on $M_1 \times M_2$. Given $R = I_1 \times I_2 \in \mathcal{M}_1(\Omega)$, let $\widehat{I_2} = \widehat{I_2}(I_1)$ be the biggest dyadic cube containing I_2 such that

$$\mu\left((I_1 \times \widehat{I_2})\cap\Omega\right) > \frac{1}{2}\mu(I_1 \times \widehat{I_2}),$$

where $\mu = \mu_1 \times \mu_2$ is the measure on $\widetilde{M}$. Similarly, Given $R = I_1 \times I_2 \in \mathcal{M}_2(\Omega)$, let $\widehat{I_1} = \widehat{I_1}(I_2)$ be the biggest dyadic cube containing I_1 such that

$$\mu\left((\widehat{I_1} \times I_2)\cap\Omega\right) > \frac{1}{2}\mu(\widehat{I_1} \times I_2).$$

For $I_i = I_{\tau_i}^{k_i} \subset M_i$, we denote by $(I_i)_k$, $k \in \mathbb{N}$, any dyadic cube $I_{\beta_i}^{k_i-k}$ containing $I_{\tau_i}^{k_i}$, and $(I_i)_0 = I_i$, where $i = 1, 2$. Moreover, let $w(x)$ be any increasing function such that $\sum_{j=0}^{\infty} j w(C_0 2^{-j}) < \infty$, where C_0 is any given positive constant. In applications, we may take $w(x) = x^\delta$ for any $\delta > 0$.

The Journé-type covering lemma on $\widetilde{M}$ is the following:

Lemma 2.2. *Let Ω be any open subset in $\widetilde{M}$ with finite measure. Then there exists a positive constant C such that*

$$\sum_{R=I_1\times I_2\in\mathcal{M}_1(\Omega)} \mu(R) w\left(\frac{\ell(I_2)}{\ell(\widehat{I_2})}\right) \le C\mu(\Omega) \tag{2.1}$$

and

$$\sum_{R=I_1\times I_2\in\mathcal{M}_2(\Omega)} \mu(R) w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right) \le C\mu(\Omega). \tag{2.2}$$

Proof. It suffices to prove (2.2) since (2.1) follows similarly. Following [25], let $R = I_1 \times I_2 \in \mathcal{M}_2(\Omega)$ and for $k \in \mathbb{N}$ let

$$A_{I_1,k} = \cup\left\{I_2 : I_1 \times I_2 \in \mathcal{M}_2(\Omega) \text{ and } \widehat{I_1} = (I_1)_{k-1}\right\}$$

where we use $(I_1)_1$ to denote the father of I_1 in the setting of dyadic cubes in M_1. Hence, $(I_1)_{k-1}$ means the ancestor of I_1 at $(k-1)$-level. We also denote the set

$$A(\Omega) = \{I_1 \subset M_1 : \text{ dyadic, and } \exists \text{ a dyadic } I_2 \in M_2, \text{ such that } I_1 \times I_2 \in \mathcal{M}_2(\Omega)\}.$$

We rewrite the left side in (2.2) as

$$\sum_{R=I_1\times I_2\in\mathcal{M}_2(\Omega)} \mu(R) w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right) = \sum_{I_1\in A(\Omega)} \mu_1(I_1) \sum_{k=1}^{\infty} \sum_{I_2:\ I_2\in A_{I_1,k}} \mu_2(I_2) w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right).$$

Note that from the definition of $A_{I_1,k}$, we have that for $k \in \mathbb{N}$ and $\widehat{I_1} = (I_1)_{k-1}$, $\frac{\ell(I_1)}{\ell(\widehat{I_1})} \le C2^{-k}$. This yields

$$\begin{aligned}\sum_{R=I_1\times I_2\in\mathcal{M}_2(\Omega)} \mu(R) w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right) &\le \sum_{I_1\in A(\Omega)} \mu_1(I_1) \sum_{k=1}^{\infty} w(C2^{-k}) \sum_{I_2:\ I_2\in A_{I_1,k}} \mu_2(I_2) \\ &\le \sum_{I_1\in A(\Omega)} \mu_1(I_1) \sum_{k=1}^{\infty} w(C2^{-k}) \mu_2(A_{I_1,k}),\end{aligned} \tag{2.3}$$

where the second inequality follows from the fact that all I_2 in $A_{I_1,k}$ are disjoint since I_2 are the maximal dyadic cubes and $\widehat{I_1} = (I_1)_{k-1}$ for each fixed $k \in \mathbb{N}$. We now estimate $\mu_2(A_{I_1,k})$. For any $x_2 \in A_{I_1,k}$, by the definition of $A_{I_1,k}$, there exists some dyadic cube I_2 such that $I_1 \times I_2 \in \mathcal{M}_2(\Omega)$, $x_2 \in I_2$, and $\widehat{I_1} = (I_1)_{k-1}$ for some $k \in \mathbb{N}$. Thus, by the definition of $\widehat{I_1}$, $\mu\big((I_1)_{k-1} \times I_2 \cap \Omega\big) > \frac{1}{2}\mu\big((I_1)_{k-1} \times I_2\big)$ and $\mu\big((I_1)_k \times I_2 \cap \Omega\big) \le \frac{1}{2}\mu\big((I_1)_k \times I_2\big)$. Now set $E_{I_1}(\Omega) = \cup\{I_2 : I_1 \times I_2 \subset \Omega\}$, then from the last inequality above, we have

$$\mu\big((I_1)_k \times (I_2 \cap E_{(I_1)_k})\big) \le \frac{1}{2}\mu\big((I_1)_k \times I_2\big),$$

which implies that $\mu_2(I_2 \cap E_{(I_1)_k}) \le \frac{1}{2}\mu_2(I_2)$ and hence $\mu_2(I_2 \cap (E_{(I_1)_k})^c) > \frac{1}{2}\mu_2(I_2)$, where we denote $(E_{(I_1)_k})^c = E_{I_1} \backslash E_{(I_1)_k}$. This gives

$$M_{HL,2}\big(\chi_{E_{I_1}\backslash E_{(I_1)_k}}\big)(x_2) > \frac{1}{2},$$

and hence $A_{I_1,k} \subset \big\{x_2 \in M_2 : M_{HL,2}\big(\chi_{E_{I_1}\backslash E_{(I_1)_k}}\big)(x_2) > \frac{1}{2}\big\}$, which implies that

$$\begin{aligned}\mu_2(A_{I_1,k}) &\le \mu_2\left(\left\{x_2 \in M_2 : M_{HL,2}\left(\chi_{E_{I_1}\backslash E_{(I_1)_k}}\right)(x_2) > \frac{1}{2}\right\}\right) \\ &\le C\mu_2(E_{I_1}\backslash E_{(I_1)_k}),\end{aligned} \tag{2.4}$$

where we use $M_{HL,2}$ to denote the Hardy–Littlewood maximal function on M_2.

Thus, combining the estimates of (2.3) and (2.4), we obtain

$$\sum_{R=I_1\times I_2\in\mathcal{M}_2(\Omega)} \mu(R) w\left(\frac{\ell_1(I_1)}{\ell_1(\hat{I}_1)}\right) \le C \sum_{I_1\in A(\Omega)} \mu_1(I_1) \sum_{k=1}^{\infty} w(C2^{-\kappa_1 k}) \mu_2(E_{I_1}\backslash E_{(I_1)_k}).$$

Next, we point out that for each $k \in \mathbb{N}$,

$$\begin{aligned}\mu_2(E_{I_1}\backslash E_{(I_1)_k}) &\le \mu_2(E_{I_1}\backslash E_{(I_1)_1}) + \cdots + \mu_2(E_{(I_1)_{k-1}}\backslash E_{(I_1)_k}) \\ &\le C \sum_{\substack{\widetilde{I} \text{ dyadic} \\ I_1\subseteq\widetilde{I}\subsetneqq(I_1)_k,\ \widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})\subset\Omega}} \mu_2(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1}),\end{aligned}$$

where the last inequality follows from the definition of $(I_1)_k$. As a consequence,

$$\sum_{R=I_1\times I_2\in\mathcal{M}_2(\Omega)} \mu(R) w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)$$
$$\leq C \sum_{I_1\in A(\Omega)} \mu_1(I_1) \sum_{k=1}^{\infty} w(C2^{-k}) \sum_{\substack{\widetilde{I}:\,\text{dyadic}\\ I_1\subseteq\widetilde{I}\subsetneqq(I_1)_k,\ \widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})\subset\Omega}} \mu_2(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1}).$$

Now interchanging the order of the sums we can obtain that the above inequality is bounded by

$$C\sum_{k=1}^{\infty} w(C2^{-k}) \sum_{\substack{\widetilde{I}:\,\text{dyadic}\\ \widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})\subset\Omega}} \mu_1(\widetilde{I})\mu_2(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1}) \sum_{\substack{I_1:\,\text{dyadic}\\ I_1\subseteq\widetilde{I}\subsetneqq(I_1)_k,}} \frac{\mu_1(I_1)}{\mu_1(\widetilde{I})}$$
$$\leq C\sum_{k=1}^{\infty} w(C2^{-\kappa_1 k}) \sum_{\substack{\widetilde{I}:\,\text{dyadic}\\ \widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})\subset\Omega}} \mu_1(\widetilde{I})\mu_2(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1}) \sum_{j=1}^{k} \sum_{\substack{I_1:\,\text{dyadic}\\ I_1\subseteq\widetilde{I}\subsetneqq(I_1)_j,}} \frac{\mu_1(I_1)}{\mu_1(\widetilde{I})}.$$

Note that for each $j = 1, \ldots, k$, $\sum_{\substack{I_1:\,\text{dyadic}\\ I_1\subseteq\widetilde{I}\subsetneqq(I_1)_j,}} \frac{\mu_1(I_1)}{\mu_1(\widetilde{I})} \leq C$, where C is a constant independent of I_1, $\widetilde{I}$. Hence,

$$\sum_{R=I_1\times I_2\in\mathcal{M}_2(\Omega)} \mu(R) w\left(\frac{\ell_1(I_1)}{\ell_1(\hat{I}_1)}\right) \leq C\sum_{k=1}^{\infty} k w(C2^{-\kappa_1 k}) \sum_{\substack{\widetilde{I}:\,\text{dyadic}\\ \widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})\subset\Omega}} \mu_1(\widetilde{I})\mu_2(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})$$
$$\leq C\sum_{k=1}^{\infty} k w(C2^{-\kappa_1 k})\mu(\Omega) \leq C\mu(\Omega),$$

since $\widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})$ are contained in $\{\widetilde{I}$ dyadic, $\widetilde{I}\times(E_{\widetilde{I}}\backslash E_{(\widetilde{I})_1})\subset\Omega\}$ and are disjoint. □

The proof of Lemma 2.2 is concluded. This covering lemma will be a key tool to obtain an atomic decomposition for $H^p(\widetilde{M})$, which will be given in next subsection.

2.1.2. *Atomic decomposition*

In this subsection we will apply Journé-type covering lemma to provide an atomic decomposition for $H^p(\widetilde{M})$. We point out that the atomic decomposition provided in this subsection is different from the classical ones. More precisely, we will prove an atomic decomposition for $L^q(\widetilde{M})\cap H^p(\widetilde{M})$ for any $1 < q < \infty$, where the decomposition converges in both $L^q(\widetilde{M})$ and $H^p(\widetilde{M})$ norms. In particular, the convergence in both $L^2(\widetilde{M})$ and $H^p(\widetilde{M})$ norms will be crucial for proving the boundedness of Calderón–Zygmund operators from $H^p(\widetilde{M})$ to $L^p(\widetilde{M})$.

Suppose that $\max\left(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\right) < p \le 1$ and $1 < q < \infty$. We first define a (p, q)-atom for the Hardy space $H^p(\widetilde{M})$ as follows.

Definition 2.3. A function $a(x_1,x_2)$ defined on $\widetilde{M}$ is called a (p,q)-atom of $H^p(\widetilde{M})$ if $a(x_1, x_2)$ satisfies:

(1) $\operatorname{supp} a \subset \Omega$, where Ω is an open set of $\widetilde{M}$ with finite measure;
(2) $\|a\|_{L^q} \le \mu(\Omega)^{1/q-1/p}$;
(3) a can be further decomposed into rectangle (p, q)-atoms a_R associated to dyadic rectangle $R = I_1 \times I_2$, satisfying the following
   (i) there exist two constants C_1 and C_2 such that $\operatorname{supp} a_R \subset C_1 I_1 \times C_2 I_2$;
   (ii) $\int_{M_1} a_R(x_1, x_2)dx_1 = 0$ for a.e. $x_2 \in M_2$ and $\int_{M_2} a_R(x_1, x_2)dx_2 = 0$ for a.e. $x_1 \in M_1$;
   (iii-a) for $2 \le q < \infty$, $a = \sum_{R\in\mathcal{M}(\Omega)} a_R$ and $\left(\sum_{R\in\mathcal{M}(\Omega)} \|a_R\|_{L^q}^q\right)^{1/q} \le \mu(\Omega)^{1/q-1/p}$.
   (iii-b) for $1 < q < 2$, $a = \sum_{R\in\mathcal{M}_1(\Omega)} a_R + \sum_{R\in\mathcal{M}_2(\Omega)} a_R$ and for some $\delta > 0$, there exists a constant $C_{q,\delta}$ such that

$$\left(\sum_{R\in\mathcal{M}_1(\Omega)} \left(\frac{\mu_2(I_2)}{\mu_2(\widehat{I_2})}\right)^{\delta} \|a_R\|_{L^q}^q + \sum_{R\in\mathcal{M}_2(\Omega)} \left(\frac{\mu_1(I_1)}{\mu_1(\widehat{I_1})}\right)^{\delta} \|a_R\|_{L^q}^q\right)^{1/q} \le C_{q,\delta}\mu(\Omega)^{1/q-1/p}.$$

We remark that, when $\widetilde{M} = \mathbb{R}^n \times \mathbb{R}^m$, a $(p, 2)$-atom with the conditions (i), (ii) and (iii-a) ($q = 2$) was introduced by R. Fefferman [9]. Note that the condition in (iii-b) is new.

The main result in this subsection is the following:

Theorem 2.4. *Suppose that* $\max\left(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\right) < p \le 1 < q < \infty$. *Then* $f \in L^q(\widetilde{M}) \cap H^p(\widetilde{M})$ *if and only if f has an atomic decomposition; that is,*

$$f = \sum_{i=-\infty}^{\infty} \lambda_i a_i,$$

where a_i are (p, q)-atoms, $\sum_i |\lambda_i|^p < \infty$, and the series converges in both $H^p(\widetilde{M})$ and $L^q(\widetilde{M})$. Moreover,

$$\|f\|_{H^p(\widetilde{M})} \approx \inf\left\{\left\{\sum_i |\lambda_i|^p\right\}^{\frac{1}{p}}, f = \sum_i \lambda_i a_i\right\},$$

where the infimum is taken over all decompositions as above and the implicit constants are independent of the $L^q(\widetilde{M})$ and $H^p(\widetilde{M})$ norms of f.

Proof of Theorem 2.4. Let $f \in L^q(\widetilde{M}) \cap H^p(\widetilde{M})$. We prove that f has an atomic decomposition. The key tool to do this is the following Calderón reproducing formula in (1.5).

$$f(x_1,x_2) = \sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{I_1}\sum_{I_2}\mu_1(I_1)\mu_2(I_2) \times D^1_{k_1}(x_1,x_{I_1})D^2_{k_2}(x_2,x_{I_2})\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1},x_{I_2}) \tag{2.5}$$

where the series converges in the norm of $L^q(\widetilde{M})$, $1<q<\infty$ and $H^p(\widetilde{M})$. Note that as a function of x_1, $D^1_{k_1}(x_1,x_{I_1})$ is supported in $\{x_1 : d_1(x_1,x_{I_1}) \le C2^{-k_1+N_1}\}$ and similarly for $D^2_{k_2}(x_2,x_{I_2})$. For each $k\in\mathbb{Z}$, let

$$\Omega_k = \{(x_1,x_2)\in M_1\times M_2 : \widetilde{\widetilde{S}}(f)(x_1,x_2) > 2^k\},$$

where $\widetilde{\widetilde{S}}(f)$ is similar to $S(f)$ but with $D^1_{k_1}D^2_{k_2}$ replaced by $\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}$. More precisely,

$$\widetilde{\widetilde{S}}(f)(x_1,x_2) = \left\{\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{I_1}\sum_{I_2}|\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_1,x_2)|^2\chi_{I_1}(x_1)\chi_{I_2}(x_2)\right\}^{1/2}.$$

By a result in [14], $\|S(f)\|_p \approx \|\widetilde{\widetilde{S}}(f)\|_p$ for $\max\left(\frac{\omega_1}{\omega_1+\vartheta_1},\frac{\omega_2}{\omega_2+\vartheta_2}\right) < p < \infty$ and therefore, $\|f\|_{H^p(\widetilde{M})} \approx \|\widetilde{\widetilde{S}}(f)\|_p$.

Set $\widetilde{\Omega}_k = \{(x_1,x_2)\in M_1\times M_2 : \mathcal{M}_s(\chi_{\Omega_k})(x_1,x_2) > \widetilde{C}\}$, where $\mathcal{M}_s$ is the strong maximal function on $\widetilde{M}$ and $\widetilde{C}$ is a constant to be decided later. Let

$$B_k = \left\{R = I_1\times I_2 : \mu(\Omega_k\cap R) > \frac{1}{2}\mu(R),\ \text{and}\ \mu(\Omega_{k+1}\cap R) \le \frac{1}{2}\mu(R)\right\}.$$

Rewrite the summation $\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{I_1}\sum_{I_2}$ in (2.5) as $\sum_{k=-\infty}^{\infty}\sum_{R=I_1\times I_2\in B_k}$. Then we have

$$f(x_1,x_2) = \sum_{k=-\infty}^{\infty}\lambda_k a_k(x_1,x_2),$$

where

$$a_k(x_1,x_2) = \frac{1}{\lambda_k}\sum_{R=I_1\times I_2\in B_k}\mu_1(I_1)\mu_2(I_2)D^1_{k_1}(x_1,x_{I_1})D^2_{k_2}(x_2,x_{I_2})\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1},x_{I_2})$$

and

$$\lambda_k = C\left\|\left\{\sum_{R=I_1\times I_2\in B_k}\left|\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1},x_{I_2})\chi_R(\cdot,\cdot)\right|^2\right\}^{1/2}\right\|_q \mu(\widetilde{\Omega}_k)^{1/p-1/q}$$

when $2 \le q < \infty$, and for $1 < q < 2$,

$$\lambda_k = C \Big\| \Big\{ \sum_{R=I_1\times I_2 \in B_k} \Big| \widetilde{\widetilde{D}}^1{}_{k_1} \widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1}, x_{I_2}) \chi_R(\cdot,\cdot) \Big|^2 \Big\}^{1/2} \Big\|_2 \mu(\widetilde{\Omega}_k)^{1/p-1/2}.$$

Next, using the duality argument we obtain that $\| \sum_{|k|>\ell} \lambda_k a_k(x_1,x_2)\|_q \to 0$ as $\ell \to \infty$, which yields that the atomic decomposition $\sum_{k=-\infty}^{\infty} \lambda_k a_k(x_1,x_2)$ converges to f in the L^q norm.

To see that a_k has compact support, by choosing $\widetilde{C}$ sufficiently small, we can conclude that $\operatorname{supp} a_k \subset \widetilde{\Omega}_k$ since $D^1_{k_1}(x_1, x_{I_1})$ and $D^2_{k_2}(x_2, x_{I_2})$, as functions of x_1 and x_2, respectively, have compact supports with diameters being equivalent to 2^{-k_1} and 2^{-k_2}, respectively. This implies that a_k satisfies the condition (1) of Definition 2.3.

We now verify that a_k satisfies (2) of Definition 2.3. By the duality argument, we have

$$\begin{aligned} &\Big\| \sum_{R=I_1\times I_2\in B_k} \mu(R) D^1_{k_1}(\cdot, x_{I_1}) D^2_{k_2}(\cdot, x_{I_2}) \widetilde{\widetilde{D}}^1{}_{k_1} \widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1}, x_{I_2}) \Big\|_q \\ &\qquad \le C \Big\| \Big\{ \sum_{R=I_1\times I_2\in B_k} \big| \widetilde{\widetilde{D}}_{k_1} \widetilde{\widetilde{D}}_{k_2}(f)(x_{I_1}, x_{I_2}) \big|^2 \chi_R(\cdot,\cdot) \Big\}^{1/2} \Big\|_q. \end{aligned}$$

This yields that when $2 \le q < \infty$, $\|a_k\|_q \le \mu\big(\widetilde{\Omega}_k\big)^{1/q-1/p}$. And for $1 < q < 2$, since a_k is supported in $\widetilde{\Omega}_k$, applying Hölder's inequality yields

$$\|a_k\|_q \le \|a_k\|_2 \mu(\widetilde{\Omega}_k)^{1/q-1/2} \le C\mu\big(\widetilde{\Omega}_k\big)^{1/q-1/p}.$$

As a consequence, we get that a_k satisfies (2) of Definition 2.3.

It remains to check that a_k satisfies the condition (3) of Definition 2.3. To see this, we can further decompose a_k as $a_k = \sum_{\overline{R}\in\mathcal{M}(\widetilde{\Omega}_k)} a_{k,\overline{R}}$, where

$$\begin{aligned} a_{k,\overline{R}}(x_1,x_2) = \frac{1}{\lambda_k} &\sum_{R=I_1\times I_2\in B_k,\ R\subset \overline{R}} \mu_1(I_1)\mu_2(I_2) D^1_{k_1}(x_1, x_{I_1}) D^2_{k_2}(x_2, x_{I_2}) \\ &\times \widetilde{\widetilde{D}}^1{}_{k_1} \widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1}, x_{I_2}). \end{aligned}$$

Similar to a_k, we can verify that $\operatorname{supp} a_{k,\overline{R}} \subset C\overline{R}$ and by the facts that

$$\int D^1_{k_1}(x_1, x_{I_1}) d\mu^1(x_1) = \int D^2_{k_2}(x_2, x_{I_2}) d\mu^2(x_2) = 0,$$

$$\int_{M_1} a_{k,\overline{R}}(x_1,x_2) d\mu^1(x_1) = 0 \quad \text{for a.e.} \quad x_2 \in M_2,$$

and

$$\int_{M_2} a_{k,\overline{R}}(x_1, x_2)d\mu^2(x_2) = 0 \quad \text{for a.e.} \quad x_1 \in M_1,$$

which yield that the conditions (i) and (ii) of (3) in Definition 2.3 hold. Now it's left to show that a_k satisfies the conditions (iii-a) and (iii-b) of (3).

For $2 \le q < \infty$, applying the same argument for the estimates of $\|a_k\|_q$ with $2 \le q < \infty$ yields

$$\left\{ \sum_{\overline{R} \in \mathcal{M}(\widetilde{\Omega}_k)} \|a_{k,\overline{R}}\|_{L^q}^q \right\}^{1/q} \le \mu(\widetilde{\Omega}_k)^{1/q-1/p},$$

which concludes that the condition (iii-a) holds. For $1 < q < 2$, by applying Hölder's inequality and Journé-type covering lemma with $\delta' = \frac{2\delta}{2-q}$, we can get that (iii-b) holds. This implies that we obtain a desired atomic decomposition for f.

To prove the converse, it suffices to verify that there is a positive constant C such that

$$\|S(a)\|_{L^p(\widetilde{M})} \le C \tag{2.6}$$

for each (p, q)-atom a of $H^p(\widetilde{M})$ with $1 < q < \infty$.

To this end, fix an (p, q)-atom a with $\operatorname{supp} a \subset \Omega$ and $a = \sum_{R \in \mathcal{M}(\Omega)} a_R$. Set $\widetilde{\Omega} = \{(x_1, x_2) \in \widetilde{M} : \mathcal{M}_s(\chi_\Omega)(x_1, x_2) > 1/2\}$ and $\widetilde{\widetilde{\Omega}} = \{(x_1, x_2) \in \widetilde{M} : \mathcal{M}_s(\chi_{\widetilde{\Omega}})(x_1, x_2) > 1/2\}$. Moreover, for any $R = I_1 \times I_2 \in \mathcal{M}_1(\Omega)$, set $\widehat{R} = \widehat{I_1} \times I_2 \subset \mathcal{M}_1(\widetilde{\Omega})$. Then $\mu(\widehat{R} \cap \Omega) > \frac{\mu(\widehat{R})}{2}$. Similarly, set $\widehat{\widehat{R}} = \widehat{I_1} \times \widehat{I_2} \subset \mathcal{M}_2(\widetilde{\widetilde{\Omega}})$. Then $\mu(\widehat{\widehat{R}} \cap \widetilde{\Omega}) > \frac{\mu(\widehat{\widehat{R}})}{2}$.

Now let $\overline{C}$ be a constant to be chosen later. We decompose $\|S(a)\|_{L^p(\widetilde{M})}^p$ as

$$\begin{aligned} &\int_{\cup_{R \in \mathcal{M}(\Omega)} 100\overline{C}\widehat{\widehat{R}}} S(a)(x_1, x_2)^p d\mu^1(x_1) d\mu^2(x_2) \\ &+ \int_{(\cup_{R \in \mathcal{M}(\Omega)} 100\overline{C}\widehat{\widehat{R}})^c} S(a)(x_1, x_2)^p d\mu^1(x_1) d\mu^2(x_2) := A + B. \end{aligned}$$

Applying Hölder's inequality, the estimate of A then follows from the L^2 boundedness of $\widetilde{S}$ and the L^2 norm of the atom a as in (2) of Definition 2.3. Using the decomposition of a as in (3) of Definition 2.3 and the fact $p \le 1$, B is bounded by

$$\sum_{R \in \mathcal{M}(\Omega)} \int_{(100\overline{C}\widehat{\widehat{R}})^c} S(a_R)(x_1, x_2)^p d\mu^1(x_1) d\mu^2(x_2).$$

Then we split $\left(100\overline{C}\widehat{\widehat{R}}\right)^c$ into two parts $\left(100\overline{C}\widehat{I_1}\right)^c \times M_2$ and $M_1 \times \left(100\overline{C}\widehat{I_2}\right)^c$ and denote these two parts by B_1 and B_2, respectively. It suffices to estimate B_1 since the estimate for B_2 is similar by the symmetry. To estimate B_1, we continue to split it into two cases $\big(100\overline{C}\widehat{I_1}\big)^c \times 100\overline{C}I_2$ and $\big(100\overline{C}\widehat{I_1}\big)^c \times (100\overline{C}I_2)^c$ and denote these two cases by B_{11} and B_{12}. Applying Hölder's inequality on the second variable and the vector-valued Littlewood-Paley estimate, B_{11} is bounded by

$$C \sum_{R\in\mathcal{M}(\Omega)} \mu_2(I_2)^{1-p/q} \int_{x_1\notin 100\overline{C}\widehat{I_1}} \Big[\int_{M_2} \Big[\sum_{k_1=-\infty}^{\infty} \Big|\int_{M_1} D^1_{k_1}(x_1,y_1) \times a_R(y_1,x_2)d\mu^1(y_1)\Big|^2\Big]^{\frac{q}{2}} d\mu^2(x_2)\Big]^{p/q} d\mu^1(x_1).$$

Using the cancellation condition of the atom a_R and writing $\big|\int_{M_1} D^1_{k_1}(x_1,y_1) \times a_R(y_1,x_2)d\mu^1(y_1)\big| = \big|\int_{M_1}[D^1_{k_1}(x_1,y_1) - D^1_{k_1}(x_1,z_1)]a_R(y_1,x_2)d\mu^1(y_1)\big|$, where z_1 is the center of I_1, and then applying the smoothness conditions on $D^1_{k_1}$ imply that $\big|\int_{M_1} D^1_{k_1}(x_1,y_1)a_R(y_1,x_2)d\mu^1(y_1)\big|$ is bounded by

$$C2^{k_1\vartheta_1}\ell(I_1)^{\vartheta_1}\left(\frac{1}{V_{2^{-k_1}}(x_1)+V_{2^{-k_1}}(z_1)+V(x_1,z_1)}\right)\int_{M_1}|a_R(y_1,x_2)|d\mu^1(y_1).$$

Inserting this estimate back yields that B_{11} is dominated by

$$C\sum_{R\in\mathcal{M}(\Omega)} \mu(R)^{1-p/q}\|a_R\|^p_{L^q(\widetilde{M})}\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)^{p\vartheta_1}\left(\frac{V(z_1,\ell(\widehat{I_1}))}{\mu_1(I_1)}\right)^{1-p}.$$

Note that $\left(\frac{V(z_1,\ell(\widehat{I_1}))}{\mu_1(I_1)}\right)^{1-p} \le C\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)^{Q_1(1-p)}$. The above quantity is bounded by

$$C\sum_{R\in\mathcal{M}(\Omega)} \mu(R)^{1-p/q}\|a_R\|^p_{L^q(\widetilde{M})}\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)^{p\vartheta_1-Q_1(1-p)}. \tag{2.7}$$

This yields that when $2\le q<\infty$,

$$B_{11}\le C\sum_{R\in\mathcal{M}(\Omega)}\|a_R\|^p_{L^q(\widetilde{M})}\mu(R)^{1-p/q}w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right),$$

where $w(x)=x^\alpha$ with $\alpha = p\vartheta_1 - Q_1(1-p)$. Note that $\alpha>0$ since $p>\frac{Q\omega 1}{Q\omega 1+\vartheta_1}$. Then, applying Hölder's inequality and the Journé-type covering lemma gives

$$\begin{aligned}B_{11} &\le C\left(\sum_{R\in\mathcal{M}(\Omega)}\|a_R\|^q_{L^q(\widetilde{M})}\right)^{p/q}\left(\sum_{R\in\mathcal{M}(\Omega)}\mu(R)w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)\right)^{1-p/q}\\ &\le C\mu(\Omega)^{p/q-1}\mu(\Omega)^{1-p/q}\le C.\end{aligned}$$

For $1 < q < 2$, setting $\overline{w} = w^{\frac{1}{2}}$, $\widetilde{w} = \overline{w}^{\frac{q}{q-p}}$ and $\widetilde{\widetilde{w}} = \overline{w}^{\frac{q}{p}}$ and applying the same estimate as above imply that

$$B_{11} \le C \sum_{R\in\mathcal{M}(\Omega)} \|a_R\|^p_{L^q(\widetilde{M})} \mu(R)^{1-p/q} w\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)$$
$$\le C \sum_{R\in\mathcal{M}(\Omega)} \|a_R\|^p_{L^q(\widetilde{M})} \overline{w}\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right) \mu(R)^{1-p/q} \overline{w}\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right).$$

Applying Hölder's inequality and Journé-type covering lemma implies

$$B_{11} \le C\left(\sum_{R\in\mathcal{M}(\Omega)} \|a_R\|^q_{L^q(\widetilde{M})} \widetilde{\widetilde{w}}\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)\right)^{p/q}\left(\sum_{R\in\mathcal{M}(\Omega)} \mu(R)\widetilde{w}\left(\frac{\ell(I_1)}{\ell(\widehat{I_1})}\right)\right)^{1-p/q}$$
$$\le C\mu(\Omega)^{p/q-1}\mu(\Omega)^{1-p/q} \le C.$$

We now estimate B_{12}. Using the cancellation condition of the atoms a_R, we write B_{12} as

$$\sum_{R\in\mathcal{M}(\Omega)} \int_{x_1\notin 100\overline{C}\widehat{I_1}} \int_{x_2\notin 100\overline{C}I_2} \left| \sum_{k_1=-\infty}^{\widehat{k_1}} \sum_{k_2=-\infty}^{\widehat{k_2}} \left| \int_{\widetilde{M}} [D^1_{k_1}(x_1,y_1) - D^1_{k_1}(x_1,z_1)] \right.\right.$$
$$\left.\left. \times [D^2_{k_2}(x_2,y_2) - D^2_{k_2}(x_2,z_2)] a_R(y_1,y_2) d\mu^1(y_1) d\mu^2(y_2)\right|^q\right|^{p/q} dx_1 dx_2,$$

where the constants $\widehat{k_1}$ and $\widehat{k_2}$ satisfy $2^{-\widehat{k_1}} \approx \ell(\widehat{I_1})$ and $2^{-\widehat{k_2}} \approx \ell(I_2)$, respectively. Applying smoothness properties of $D^1_{k_1}(x_1,y_1)$ and $D^2_{k_2}(x_2,y_2)$ yields that B_{12} satisfies the same estimate as B_{11} as in (2.7). This concludes the proof of Theorem 2.4. For more details of the proof, we refer the readers to the long version of this paper [13]. □

2.1.3. $H^p \to L^p$ *boundedness*

In this subsection applying the atomic decomposition provided in the previous subsection, we prove that if T is a product Calderón-Zygmund operator, then T can be extended to a bounded operator from $H^p(\widetilde{M})$ to $L^p(\widetilde{M})$. Note that if T is a product Calderón-Zygmund operator then $K(x_1,y_1,x_2,y_2)$, the kernel of T, satisfies the following estimates

$$\|K_i(x_i,y_i) - K_i(x_i,y_i')\|_{L^2} \le C\left(\frac{d_i(y_i,y_i')}{d_i(x_i,y_i)}\right)^{\varepsilon_i} V(x_i,y_i)^{-1}$$

if $d_i(y_i,y_i') \le d_i(x_i,y_i)/2A_i$, for $i = 1, 2$. From this it is easy to see that

$$|K(x_1,y_1,x_2,y_2) - K(x_1,y_1',x_2,y_2) - K(x_1,y_1,x_2,y_2') + K(x_1,y_1',x_2,y_2')|$$
$$\le C\left(\frac{d_1(y_1,y_1')}{d_1(x_1,y_1)}\right)^{\varepsilon_i} V(x_1,y_1)^{-1}\left(\frac{d_2(y_2,y_2')}{d_2(x_2,y_2)}\right)^{\varepsilon_i} V(x_2,y_2)^{-1}$$

if $d_1(y_1,y_1') \le d_1(x_1,y_1)/2A_1$ and $d_2(y_2,y_2') \le d_2(x_2,y_2)/2A_2$.

The main result of this subsection is the following:

Theorem 2.5. *For* $\max\left(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\right) < p \leq 1$, *$T$ extends to a bounded operator from $H^p(\widetilde{M})$ to $L^p(\widetilde{M})$. Moreover, there exists a constant C such that* $\|Tf\|_{L^p(\widetilde{M})} \leq C\|f\|_{H^p(\widetilde{M})}$.

Proof. Fix $\max\left(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\right) < p \leq 1$. Since $H^p(\widetilde{M}) \cap L^2(\widetilde{M})$ is dense in $H^p(\widetilde{M})$, it suffices to prove that there exists a positive constant C such that for every $f \in H^p(\widetilde{M}) \cap L^2(\widetilde{M})$,

$$\|Tf\|_{L^p(\widetilde{M})} \leq C\|f\|_{H^p(\widetilde{M})}. \tag{2.8}$$

The proof of the estimate of (2.8) is similar to the proof of Theorem 2.4 with $q = 2$. Indeed, we only need to show that there exists a positive constant C such that for any $(p, 2)$-atom a of $H^p(\widetilde{M})$,

$$\|Ta\|_{L^p(\widetilde{M})} \leq C.$$

And the proof of this estimate is similar to the proof of (2.6) with $q = 2$. To see this, we decompose $\|T(a)\|^p_{L^p(\widetilde{M})}$ by

$$\begin{aligned}
&\int_{\cup_{R\in\mathcal{M}(\Omega)}100\overline{C}\widehat{\widehat{R}}} T(a)(x_1,x_2)^p d\mu^1(x_1)d\mu^2(x_2)\\
&+\sum_{R\in\mathcal{M}(\Omega)}\int_{(100\overline{C}\widehat{\widehat{R}})^c} T(a_R)(x_1,x_2)^p d\mu^1(x_1)d\mu^2(x_2)\\
:=\ & A+B.
\end{aligned}$$

Furthermore, similarly decompose $B = B_1 + B_2$ and $B_1 = B_{11} + B_{12}$. Applying the similar estimates, it is easy to verify that B_{11} and B_{12} satisfy the estimate in (2.7) with $q = 2$ and hence repeating the same proof concludes the Theorem 2.5. For more details of the proof, we refer the readers to the long version of this paper [13]. □

2.1.4. $L^\infty \to$ *BMO boundedness*

As a consequence of Theorem 2.5 with $p = 1$, together with the duality that $(H^1(\widetilde{M}))^* = BMO(\widetilde{M})$, we obtain the following

Theorem 2.6. *Suppose that T is a Calderón-Zygmund operator. Then T extends to a bounded operator from $L^\infty(\widetilde{M})$ to $BMO(\widetilde{M})$. Moreover, there exists a constant C such that*

$$\|Tf\|_{BMO(\widetilde{M})} \leq C\|f\|_\infty.$$

We point out that Theorem 2.6 gives the necessary condition of Theorem 1 as follows.

Corollary 2.7. *Suppose that T and $\widetilde{T}$ are Calderón-Zygmund operators. Then $T(1)$, $T^*(1)$, $\widetilde{T}(1)$ and $(\widetilde{T})^*(1)$ lie in $BMO(\widetilde{M})$.*

Proof of Theorem 2.6. Suppose that T is a Calderón-Zygmund operator defined in Subsection 3.1. We have to define Tf for $f \in L^\infty(\widetilde{M})$. To this end, we first observe that if $f \in L^\infty(\widetilde{M}) \cap L^2(\widetilde{M})$ then Tf is well defined, and moreover, for $g \in H^1(\widetilde{M}) \cap L^2(\widetilde{M})$, $\langle Tf, g\rangle = \langle f, T^*g\rangle$. This together with the fact that, by Theorem 2.5, T^* is bounded from $H^1(\widetilde{M})$ to $L^1(\widetilde{M})$ and the duality arguments (L^1, L^∞) and (H^1, BMO) gives $Tf \in BMO(\widetilde{M})$ since $T^*g \in L^1(\widetilde{M})$ and $H^1(\widetilde{M}) \cap L^2(\widetilde{M})$ is dense in $H^1(\widetilde{M})$. To define Tf for $f \in L^\infty$, we define functions $f_j(x, y)$ by $f_j(x, y) = f(x, y)$, when $d(x, x_0) \le j, d(y, y_0) \le j$ and $f_j(x, y) = 0$, otherwise, where $x_0 \in M_1$ and $y_0 \in M_2$ are any fixed points. Then $f_j \in L^\infty(\widetilde{M}) \cap L^2(\widetilde{M})$ and thus for $g \in H^1(\widetilde{M}) \cap L^2(\widetilde{M})$,

$$\langle Tf_j, g\rangle = \langle f_j, T^*g\rangle \to \langle f, T^*g\rangle.$$

Indeed, $\|f_j\|_{L^\infty(\widetilde{M})} \le \|f\|_{L^\infty(\widetilde{M})}$, $f_j \to f$ almost everywhere, and $T^*g \in L^1(\widetilde{M})$, so that we can apply Lebesgue's dominated convergence theorem. This implies that functions Tf_j form a bounded sequence in $BMO(\widetilde{M})$ and this sequence converges to Tf in the topology (H^1, BMO). It remains to show the estimate in Theorem 2.6. To do this, we first consider $f \in L^2(\widetilde{M}) \cap L^\infty(\widetilde{M})$. Then for $g \in H^1(\widetilde{M}) \cap L^2(\widetilde{M})$, as mentioned,

$$|\langle Tf, g\rangle| \le C\|f\|_{L^\infty(\widetilde{M})}\|g\|_{H^1(\widetilde{M})}.$$

This together with the fact that $H^1(\widetilde{M}) \cap L^2(\widetilde{M})$ is dense in $H^1(\widetilde{M})$ implies that $\langle Tf, g\rangle$ defines a continuous linear functional on $H^1(\widetilde{M})$ and its norm is dominated by $C\|f\|_{L^\infty(\widetilde{M})}$. By the duality argument between $H^1(\widetilde{M})$ and $CMO^1(\widetilde{M})$, these exists $h \in CMO^1(\widetilde{M})$ such that $\langle Tf, g\rangle = \langle h, g\rangle$ for all $g \in \overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ and $\|h\|_{CMO^1(\widetilde{M})} \le C\|f\|_{L^\infty(\widetilde{M})}$. Now we point out that as a function of (y_1, y_2) $D^2_{k_2}D^1_{k_1}(x_1, y_1, x_2, y_2) \in \overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$. Taking $g(x_1, x_2) = D^2_{k_2}D^1_{k_1}(x_1, y_1, x_2, y_2)$ in the above equality yields that $D^2_{k_2}D^1_{k_1}(Tf)(x_1, x_2) = D^2_{k_2}D^1_{k_1}(h)(x_1, x_2)$ and hence for $f \in L^2(\widetilde{M}) \cap L^\infty(\widetilde{M})$,

$$\|Tf\|_{CMO^1(\widetilde{M})} = \|h\|_{CMO^1(\widetilde{M})} \le C\|f\|_{L^\infty(\widetilde{M})}.$$

For $f \in L^\infty$, by the definition for Tf, we have

$$D^2_{k_2}D^1_{k_1}(Tf)(x_1, x_2) = D^2_{k_2}D^1_{k_1}(\lim_j Tf_j)(x_1, x_2)$$

since $D^2_{k_2} D^1_{k_1}(x_1, x_2) \in \overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ so $D^2_{k_2} D^1_{k_1}(x_1, x_2) \in H^1(\widetilde{M}) \cap L^2(\widetilde{M})$. Thus

$$\begin{aligned}\|Tf\|_{CMO^1(\widetilde{M})} &= \|\lim_j Tf_j\|_{CMO^1(\widetilde{M})} \le \liminf_j \|Tf_j\|_{CMO^1(\widetilde{M})}\\ &\le C\liminf_j \|f_j\|_{L^\infty(\widetilde{M})} \le C\|f\|_{L^\infty(\widetilde{M})}.\end{aligned}$$

Note that $CMO^1(\widetilde{M}) = BMO(\widetilde{M})$. The proof of Theorem 2.6 is concluded. □

2.1.5. L^p, $1 < p < \infty$, *boundedness*

In this subsection we prove the L^p, $1<p<\infty$, boundedness, namely the following:

Theorem 2.8. *Suppose T is a Calderón-Zygmund operator. Then T extends to a bounded operator from L^p, $1 < p < \infty$, to itself. Moreover, there exists a constant C such that*

$$\|Tf\|_p \le C\|f\|_p.$$

Indeed, in [14] the following Calderón-Zygmund decomposition was obtained.

Theorem 2.9. *Let* $\max\big(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\big) < p_2 < p < p_1 < \infty$, $\alpha > 0$ *be given and* $f \in H^p(\widetilde{M})$. *Then we may write* $f = g + b$ *where* $g \in H^{p_1}(\widetilde{M})$ *and* $b \in H^{p_2}(\widetilde{M})$ *such that* $\|g\|^{p_1}_{H^{p_1}(\widetilde{M})} \le C\alpha^{p_1-p}\|f\|^p_{H^p(\widetilde{M})}$ *and* $\|b\|^{p_2}_{H^{p_2}(\widetilde{M})} \le C\alpha^{p_2-p}\|f\|^p_{H^p(\widetilde{M})}$, *where C is an absolute constant.*

As a consequence of Theorem 2.9, the following interpolation theorem was proved in [14].

Theorem 2.10. *Let* $\max\big(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}\big) < p_2 < p_1 < \infty$ *and T be a linear operator which is bounded from* $H^{p_2}(\widetilde{M})$ *to* $L^{p_2}(\widetilde{M})$ *and from* $H^{p_1}(\widetilde{M})$ *to* $L^{p_1}(\widetilde{M})$, *then T is bounded on* $H^p(\widetilde{M})$ *for* $p_2 < p < p_1$.

Note that $H^p(\widetilde{M}) = L^p(\widetilde{M})$ for $1 < p < \infty$. Now the proof of Theorem 2.8 with $1 < p < 2$ follows from Theorem 2.5 and 2.10 directly by taking $p_2 = 1$ and $p_1 = 2$. The duality argument gives the proof of Theorem 2.8 for $2 < p < \infty$.

2.2. Sufficiency of Theorem 1

In this section, we prove the sufficiency of Theorem 1. It suffices to prove that for $f, g \in \overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ with compact supports, there exists a constant C such that $|\langle g, Tf\rangle| \le C\|f\|_2\|g\|_2$. This is because, by Calderón's reproducing formula in (1.5), the collection of functions in $\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ having compact supports is dense in L^2. As described in Section 1, we write with changing the notation from I^i to I_i, $i = 1, 2$,

$$\begin{aligned}\langle g, Tf\rangle = &\sum_{k_1'}\sum_{I_1'}\sum_{k_1}\sum_{I_1}\sum_{k_2'}\sum_{I_2'}\sum_{k_2}\sum_{I_2} \mu_1(I_1')\mu^1(I_1)\mu^2(I_2')\mu^2(I_2)\\ &\times \widetilde{\widetilde{D}}{}^1_{k_1'}\widetilde{\widetilde{D}}{}^2_{k_2'}(g)(x_{I_1'}, x_{I_2'})\big\langle D^1_{k_1'}D^2_{k_2'}, TD^1_{k_1}D^2_{k_2}\big\rangle(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})\\ &\times \widetilde{\widetilde{D}}{}^1_{k_1}\widetilde{\widetilde{D}}{}^2_{k_2}(f)(x_{I_1}, x_{I_2}).\end{aligned} \tag{2.9}$$

To see the above equality, we first consider one parameter case. Let $f_1, g_1 \in \overset{\circ}{G}_\vartheta(\beta, \gamma)(M_1)$ with compact supports and T_1 be a singular integral operator on

M_1. Then by the Carlderón reproducing formula on M_1,

$$\begin{aligned}\langle g_1, T_1 f_1\rangle &= \sum_{k_1'}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})\big\langle D^1_{k_1'}(\cdot, x_{I_1'}), T_1 f_1\big\rangle \\ &= \sum_{k_1'}\sum_{I_1'}\sum_{k_1}\sum_{I_1} \mu_1(I_1')\mu_1(I_1) \\ &\quad \times \widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})\big\langle D^1_{k_1'}, T_1 D^1_{k_1}\big\rangle(x_{I_1'}, x_{I_1})\widetilde{D}^1{}_{k_1}(f_1)(x_{I_1}).\end{aligned} \tag{2.10}$$

For the equality (2.10), we use the fact that $\sum_{k_1'>0}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})D^1_{k_1'}(x_1, x_{I_1'})$ converges in the test function space $\overset{\circ}{G}_\vartheta(\beta,\gamma)(M_1)$ with compact support, so that

$$\begin{aligned}&\Big\langle \sum_{k_1'>0}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})D^1_{k_1'}(\cdot, x_{I_1'}), T_1 f_1\Big\rangle \\ &= \sum_{k_1'>0}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})\langle D^1_{k_1'}(\cdot, x_{I_1'}), T_1 f_1\rangle.\end{aligned}$$

This, however, is not true for the term $\sum_{k_1'\le 0}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})D^1_{k_1'}(x_1,x_{I_1'})$, since the support of $D^1_{k_1'}(x_1,x_{I_1'})$ gets bigger as k_1' tends to $-\infty$, even though $\sum_{k_1'\le 0}\sum_{I_1'} \mu_1(I_1')\cdot \widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})D^1_{k_1'}(x_1, x_{I_1'}) \in \overset{\circ}{G}_\vartheta(\beta,\gamma)(M_1)$ having compact support. Now if $\theta \in \overset{\circ}{G}_\vartheta(\beta,\gamma)(M_1)$ and has compact support, then $\theta(x_1)\sum_{k_1'\le 0}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})\cdot D^1_{k_1'}(x_1, x_{I_1'})$ converges in the topology of $C_0^\beta(M_1)$. If we choose $\theta = 1$ on a large enough set which contains the support of f_1, then, by the standard estimate on the kernel of T_1, $\big\langle(1-\theta)\sum_{k_1'\le 0}\sum_{I_1'} \mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})D^1_{k_1'}(\cdot, x_{I_1'}), T_1 f_1\big\rangle = \sum_{k_1'\le 0}\sum_{I_1'} \mu_1(I_1')\cdot\widetilde{\widetilde{D}}^1{}_{k_1'}(g)(x_{I_1'})\langle(1-\theta)D^1_{k_1'}(\cdot, x_{I_1'}), T_1 f_1\rangle$. This implies the equality (2.10). For fixed k_1' we can do the same thing to f_1 to obtain the second equality. Repeating the same things above twice, first on M_1 and then on M_2, gives (2.9).

As described in Section 1, we decompose the bilinear form $\langle g, Tf\rangle$ as

$$\langle g, Tf\rangle = \langle g, Tf\rangle_{\text{Case 1}} + \langle g, Tf\rangle_{\text{Case 2}} + \langle g, Tf\rangle_{\text{Case 3}} + \langle g, Tf\rangle_{\text{Case 4}},$$

where case 1: $k_1^{'} \geq k_1$ and $k_2^{'} \geq k_2$; case 2: $k_1^{'} \geq k_1$ and $k_2^{'} < k_2$; case 3: $k_1^{'} < k_1$ and $k_2^{'} \geq k_2$; case 4: $k_1^{'} < k_1$ and $k_2^{'} < k_2$. More precisely,

$$\begin{aligned}\langle g, Tf\rangle_{\text{Case 1}} = & \sum_{k_1 \leq k_1^{'}} \sum_{k_2 \leq k_2^{'}} \sum_{I_1^{'}} \sum_{I_2^{'}} \sum_{I_1} \sum_{I_2} \mu_1(I_1^{'})\mu_1(I_1)\mu_2(I_2^{'})\mu_2(I_2) \\ & \times \widetilde{\widetilde{D}}^1_{k_1^{'}} \widetilde{\widetilde{D}}^2_{k_2^{'}}(g)(x_{I_1^{'}}, x_{I_2^{'}}) \\ & \times \widetilde{\widetilde{D}}^1_{k_1} \widetilde{\widetilde{D}}^2_{k_2}(f)(x_{I_1}, x_{I_2})\big\langle D^1_{k_1^{'}} D^2_{k_2^{'}}, T D^1_{k_1} D^2_{k_2}\big\rangle(x_{I_1^{'}}, x_{I_2^{'}}, x_{I_1}, x_{I_2})\end{aligned}$$

and similarly for other three terms.

Since the estimates for $\langle g, Tf\rangle_{\text{Case 1}}$ and $\langle g, Tf\rangle_{\text{Case 2}}$ are similar to $\langle g, Tf\rangle_{\text{Case 4}}$ and $\langle g, Tf\rangle_{\text{Case 3}}$, respectively, so we only prove that the first two terms are bounded by some constant times $\|f\|_2\|g\|_2$. This will conclude the proof of the sufficiency of Theorem 1.

To deal with the first term $\langle g, Tf\rangle_{\text{Case 1}}$, as mentioned in Section 1, for $k_1 \leq k_1^{'}$ and $k_2 \leq k_2^{'}$ we first decompose

$$\begin{aligned}&\big\langle D^1_{k_1^{'}} D^2_{k_2^{'}}, T D^1_{k_1} D^2_{k_2}\big\rangle(x_{I_1^{'}}, x_{I_2^{'}}, x_{I_1}, x_{I_2}) \\ &= \int D^1_{k_1^{'}}(x_{I_1^{'}}, u_1) D^2_{k_2^{'}}(x_{I_2^{'}}, u_2) K(u_1, u_2, v_1, v_2)[D^1_{k_1}(v_1, x_{I_1}) - D^1_{k_1}(x_{I_1^{'}}, x_{I_1})] \\ &\qquad \times [D^2_{k_2}(v_2, x_{I_2}) - D^2_{k_2}(x_{I_2^{'}}, x_{I_2})] d\mu^1(u_1) d\mu^2(u_2) d\mu^1(v_1) d\mu^2(v_2) \\ &\quad + \int D^1_{k_1^{'}}(x_{I_1^{'}}, u_1) D^2_{k_2^{'}}(x_{I_2^{'}}, u_2) K(u_1, u_2, v_1, v_2) D^1_{k_1}(x_{I_1^{'}}, x_{I_1}) \\ &\qquad \times D^2_{k_2}(v_2, x_{I_2}) d\mu^1(u_1) d\mu^2(u_2) d\mu^1(v_1) d\mu^2(v_2) \\ &\quad + \int D^1_{k_1^{'}}(x_{I_1^{'}}, u_1) D^2_{k_2^{'}}(x_{I_2^{'}}, u_2) K(u_1, u_2, v_1, v_2) D^1_{k_1}(v_1, x_{I_1}) \\ &\qquad \times D^2_{k_2}(x_{I_2^{'}}, x_{I_2}) d\mu^1(u_1) d\mu^2(u_2) d\mu^1(v_1) d\mu^2(v_2) \\ &\quad - \int D^1_{k_1^{'}}(x_{I_1^{'}}, u_1) D^2_{k_2^{'}}(x_{I_2^{'}}, u_2) K(u_1, u_2, v_1, v_2) D^1_{k_1}(x_{I_1^{'}}, x_{I_1}) \\ &\qquad \times D^2_{k_2}(x_{I_2^{'}}, x_{I_2}) d\mu^1(u_1) d\mu^2(u_2) d\mu^1(v_1) d\mu^2(v_2) \\ &=: I(x_{I_1^{'}}, x_{I_2^{'}}, x_{I_1}, x_{I_2}) + II(x_{I_1^{'}}, x_{I_2^{'}}, x_{I_1}, x_{I_2}) + III(x_{I_1^{'}}, x_{I_2^{'}}, x_{I_1}, x_{I_2}) \\ &\quad + IV(x_{I_1^{'}}, x_{I_2^{'}}, x_{I_1}, x_{I_2}).\end{aligned}$$

We then write

$$\langle g, Tf\rangle_{\text{Case 1}} = \langle g, Tf\rangle_{\text{Case 1.1}} + \langle g, Tf\rangle_{\text{Case 1.2}} + \langle g, Tf\rangle_{\text{Case 1.3}} + \langle g, Tf\rangle_{\text{Case 1.4}},$$

where

$$\langle g, Tf\rangle_{\text{Case 1.1}} = \sum_{k_1\le k_1'}\sum_{k_2\le k_2'}\sum_{I_1'}\sum_{I_2'}\sum_{I_1}\sum_{I_2}\mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2)$$
$$\times \widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)(x_{I_1'}, x_{I_2'})$$
$$\times \widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1}, x_{I_2})I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}).$$

The other terms $\langle g, Tf\rangle_{\text{Case 1.i}}, i = 2, 3, 4$, are defined similarly.

For the case 2 with $k_1' \ge k_1$ and $k_2' < k_2$, we similarly decompose $\langle g, Tf\rangle_{\text{Case 2}}$ as

$$\big\langle D^1_{k_1'} D^2_{k_2'}, T D^1_{k_1} D^2_{k_2}\big\rangle(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})$$
$$= \int D^1_{k_1'}(x_{I_1'}, u_1)[D^2_{k_2'}(x_{I_2'}, u_2) - D^2_{k_2'}(x_{I_2'}, x_{I_2})]$$
$$\times K(u_1, u_2, v_1, v_2)[D^1_{k_1}(v_1, x_{I_1}) - D^1_{k_1}(x_{I_1'}, x_{I_1})]$$
$$\times D^2_{k_2}(v_2, x_{I_2})d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2)$$
$$+ \int D^1_{k_1'}(x_{I_1'}, u_1)D^2_{k_2'}(x_{I_2'}, u_2)K(u_1, u_2, v_1, v_2)D^1_{k_1}(x_{I_1'}, x_{I_1})$$
$$\times D^2_{k_2}(v_2, x_{I_2})d\mu^{(}u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2)v_2$$
$$+ \int D^1_{k_1'}(x_{I_1'}, u_1)D^2_{k_2'}(x_{I_2'}, x_{I_2})K(u_1, u_2, v_1, v_2)D^1_{k_1}(v_1, x_{I_1})$$
$$\times D^2_{k_2}(v_2, x_{I_2})d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2)$$
$$- \int D^1_{k_1'}(x_{I_1'}, u_1)D^2_{k_2'}(x_{I_2'}, x_{I_2})K(u_1, u_2, v_1, v_2)D^1_{k_1}(x_{I_1'}, x_{I_1})$$
$$\times D^2_{k_2}(v_2, x_{I_2})d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2)$$
$$=: V(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}) + VI(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}) + VII(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})$$
$$+ VIII(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}),$$

and then write

$$\langle g, Tf\rangle_{\text{Case 2}} = \langle g, Tf\rangle_{\text{Case 2.1}} + \langle g, Tf\rangle_{\text{Case 2.2}} + \langle g, Tf\rangle_{\text{Case 2.3}} + \langle g, Tf\rangle_{\text{Case 2.4}},$$

where

$$\langle g, Tf\rangle_{\text{Case 2.1}} = \sum_{k_1\le k_1'}\sum_{k_2>k_2'}\sum_{I_1'}\sum_{I_2'}\sum_{I_1}\sum_{I_2}\mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2)$$
$$\times \widetilde{\widetilde{D^1}}_{k_1'}\widetilde{\widetilde{D^2}}_{k_2'}(g)(x_{I_1'}, x_{I_2'})$$
$$\times \widetilde{\widetilde{D^1}}_{k_1}\widetilde{\widetilde{D^2}}_{k_2}(f)(x_{I_1}, x_{I_2})V(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}).$$

Similarly for other terms $\langle g, Tf\rangle_{\text{Case 2.i}}$, $i = 2, 3, 4$. The details of proofs will be given in next subsections.

2.2.1. *Almost orthogonal estimate on* $\widetilde{M} = M_1 \times M_2$

in this subsection we deal with $\langle g, Tf\rangle_{\text{Case 1.1}}$ and $\langle g, Tf\rangle_{\text{Case 2.1}}$. The main method is the almost orthogonality argument on $\widetilde{M} = M_1 \times M_2$. Indeed, we will show that there exists a constant C such that for $k_1' > k_1$ and $k_2' > k_2$,

$$\begin{aligned}
&|I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})| \\
&= \Big| \int D^1_{k_1'}(x_{I_1'}, u_1)D^2_{k_2'}(x_{I_2'}, u_2)K(u_1, u_2, v_1, v_2)[D^1_{k_1}(v_1, x_{I_1}) - D^1_{k_1}(x_{I_1'}, x_{I_1})] \\
&\quad \times [D^2_{k_2}(v_2, x_{I_2}) - D^2_{k_2}(x_{I_2'}, x_{I_2})]d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2)\Big| \\
&\le C2^{(k_1-k_1')\varepsilon_1}2^{-(k_2-k_2')\varepsilon_2} \\
&\quad \times \frac{1}{V_{2^{-k_1}}(x_{I_1'}) + V_{2^{-k_1}}(x_{I_1}) + V(x_{I_1'}, x_{I_1})}\frac{2^{-k_1\varepsilon_1}}{(2^{-k_1} + d_1(x_{I_1'}, x_{I_1}))^{\varepsilon_1}} \\
&\quad \times \frac{1}{V_{2^{-k_2}}(x_{I_2'}) + V_{2^{-k_2}}(x_{I_2}) + V(x_{I_2'}, x_{I_2})}\frac{2^{-k_2\varepsilon_2}}{(2^{-k_2} + d_2(x_{I_2'}, x_{I_2}))^{\varepsilon_2}}.
\end{aligned} \tag{2.11}$$

We would like to remark that the cancellation condition on the kernel K is not required in the above almost orthogonality estimate and only size, smoothness on K and the weak boundedness property on T are needed. To show the above estimate, we first consider the one parameter case. The estimate for two parameter case will follow from the iterative methods. As mentioned in Section 1, let S be a singular integral operator associated with the kernel L defined on M_1 having the weak boundedness property. Then for $k_1 < k_1'$ there exists a positive constant C such that the following orthogonal estimate holds

$$\Big| \iint D^1_{k_1'}(x_1, u_1)L(u_1, v_1)[D^1_{k_1}(v_1, y_1) - D^1_{k_1}(x_1, y_1)]d\mu^1(u_1)d\mu^1(v_1)\Big|$$
$$\le C|L|_{CZ}2^{(k_1-k_1')\varepsilon_1}\frac{1}{V_{2^{-k_1}}(x_1) + V_{2^{-k_1}}(y_1) + V(x_1, y_1)}\frac{2^{-k_1\varepsilon_1}}{(2^{-k_1} + d_1(x_1, y_1))^{\varepsilon_1}}.$$

The proof of the above estimate is similar to the classical case. See [14] for the details of the proof in our setting. Now we turn to the proof of the estimate in (2.11). To see that this can be done by the iteration, we write

$$\int D^1_{k'_1}\left(x_{I'_1}, u_1\right) D^2_{k'_2}\left(x_{I'_2}, u_2\right) K\left(u_1, u_2, v_1, v_2\right)\left[D^1_{k_1}\left(v_1, x_{I_1}\right) - D^1_{k_1}\left(x_{I'_1}, x_{I_1}\right)\right]$$

$$\times\left[D^2_{k_2}\left(v_2, x_{I_2}\right) - D^2_{k_2}\left(x_{I'_2}, x_{I_2}\right)\right] d\mu^1(u_1) d\mu^2(u_2) d\mu^1(v_1) d\mu^2(v_2)$$

$$= \left\langle D^2_{k'_2}\left(x_{I'_2}, u_2\right), \left\langle D^1_{k'_1}\left(x_{I'_1}, \cdot\right), K_2(u_2, v_2)\left[D^1_{k_1}\left(\cdot, x_{I_1}\right) - D^1_{k_1}\left(x_{I'_1}, x_{I_1}\right)\right]\right\rangle\right.$$

$$\left.\times\,[D^2_{k_2}(v_2, x_{I_2}) - D^2_{k_2}(x_{I'_2}, x_{I_2})]\right\rangle,$$

where, by definition of the product singular integral operator given in Section 1, for fixed points $u_2, v_2 \in M_2$, $K_2(u_2, v_2)$ is a Calderón-Zygmund operator on M_1 with the operator norm $\|K_2(u_2, v_2)\|_{CZ(M_1)}$ which is a singular integral operator on M_2. By the estimate for one parameter case provided above, for $k'_1 > k_1$,

$$\left|\left\langle D^1_{k'_1}\left(x_{I'_1}, \cdot\right), K_2(u_2, v_2)\left[D^1_{k_1}\left(\cdot, x_{I_1}\right) - D^1_{k_1}\left(x_{I'_1}, x_{I_1}\right)\right]\right\rangle\right|$$

$$\leq C\|K_2(u_2, v_2)\|_{CZ(M_1)} 2^{(k_1-k'_1)\varepsilon_1}$$

$$\times \frac{1}{V_{2^{-k_1}}\left(x_{I'_1}\right) + V_{2^{-k_1}}\left(x_{I_1}\right) + V\left(x_{I'_1}, x_{I_1}\right)} \frac{2^{-k_1\varepsilon_1}}{\left(2^{-k_1} + d_1\left(x_{I'_1}, x_{I_1}\right)\right)^{\varepsilon_1}}.$$

Similarly,

$$\left|\left\langle D^1_{k'_1}\left(x_{I'_1}, \cdot\right), \left[K_2\left(u_2, v_2\right) - K_2\left(u_2, v'_2\right)\right]\left[D^1_{k_1}\left(\cdot, x_{I_1}\right) - D^1_{k_1}\left(x_{I'_1}, x_{I_1}\right)\right]\right\rangle\right|$$

$$\leq C\left\|K_2(u_2, v_2) - K_2(u_2, v'_2)]\right\|_{CZ(M_1)} 2^{(k_1-k'_1)\varepsilon_1}$$

$$\times \frac{1}{V_{2^{-k_1}}\left(x_{I'_1}\right) + V_{2^{-k_1}}\left(x_{I_1}\right) + V(x_{I'_1}, y_1)} \frac{2^{-k_1\varepsilon_1}}{\left(2^{-k_1} + d_1\left(x_{I'_1}, x_{I_1}\right)\right)^{\varepsilon_1}}$$

and the same estimate holds with interchanging u_2 and v_2.

This together with the fact that $\|K_2(u_2, v_2)\|_{CZ(M_1)}$ is a singular integral operator on M_2 having the weak boundedness property implies that $\langle D^1_{k'_1}(x_{I'_1}, \cdot)$, $K_2(u_2, v_2)[D^1_{k_1}(\cdot, x_{I_1}) - D^1_{k_1}(x_{I'_1}, x_{I_1})]\rangle$ is a singular integral on M_2 having the weak boundedness property. Moreover,

$$\left|\left\langle D^1_{k_1'}\left(x_{I_1'},\cdot\right), K_2(u_2,v_2)\left[D^1_{k_1}\left(\cdot,x_{I_1}\right)-D^1_{k_1}\left(x_{I_1'},x_{I_1}\right)\right]\right\rangle\right|_{CZ}$$

$$\leq C2^{(k_1-k_1')\varepsilon_1}\frac{1}{V_{2^{-k_1}}\left(x_{I_1'}\right)+V_{2^{-k_1}}\left(x_{I_1}\right)+V\left(x_{I_1'},x_{I_1}\right)}\frac{2^{-k_1\varepsilon_1}}{\left(2^{-k_1}+d_1\left(x_{I_1'},x_{I_1}\right)\right)^{\varepsilon_1}}.$$

Applying the estimate for one parameter case again yields that for $k_2' > k_2$,

$$\Big|\Big\langle D^2_{k_2'}\left(x_{I_2'},u_2\right),\left\langle D^1_{k_1'}\left(x_{I_1'},\cdot\right), K_2\left(u_2,v_2\right)\left[D^1_{k_1}\left(\cdot,x_{I_1}\right)-D^1_{k_1}\left(x_{I_1'},x_{I_1}\right)\right]\right\rangle$$

$$\times\left[D^2_{k_2}\left(v_2,x_{I_2}\right)-D^2_{k_2}\left(x_{I_2'},x_{I_2}\right)\right]\Big\rangle$$

$$\leq C\left|\left\langle D^1_{k_1'}\left(x_{I_1'},\cdot\right), K_2(u_2,v_2)\left[D^1_{k_1}\left(\cdot,x_{I_1}\right)-D^1_{k_1}\left(x_{I_1'},x_{I_1}\right)\right]\right\rangle\right|_{CZ}$$

$$\times 2^{(k_2-k_2')\varepsilon_2}\frac{1}{V_{2^{-k_2}}\left(x_{I_2'}\right)+V_{2^{-k_2}}(y_2)+V\left(x_{I_2'},y_2\right)}\frac{2^{-k_2\varepsilon_2}}{\left(2^{-k_2}+d_2\left(x_{I_2'},y_2\right)\right)^{\varepsilon_2}}$$

$$\leq C2^{(k_1-k_1')\varepsilon_2}2^{(k_2-k_2')\varepsilon_2}\frac{1}{V_{2^{-k_1}}(x_{I_1'})+V_{2^{-k_1}}(x_{I_1})+V(x_{I_1'},y_1)}\frac{2^{-k_1\varepsilon_1}}{(2^{-k_1}+d_1(x_{I_1'},x_{I_1}))^{\varepsilon_1}}$$

$$\times\frac{1}{V_{2^{-k_2}}\left(x_{I_2'}\right)+V_{2^{-k_2}}\left(x_{I_2}\right)+V\left(x_{I_2'},x_{I_2}\right)}\frac{2^{-k_2\varepsilon_2}}{\left(2^{-k_2}+d_2\left(x_{I_2'},x_{I_2}\right)\right)^{\varepsilon_2}},$$

which concludes the proof of (2.11).

Applying the Cauchy-Schwartz inequality implies that $|\langle g, Tf\rangle_{\text{Case 1.1}}|$ is bounded by

$$\Big\{\sum_{k_1\leq k_1'}\sum_{k_2\leq k_2'}\sum_{I_1'}\sum_{I_2'}\sum_{I_1}\sum_{I_2}\mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2)$$

$$\cdot\left|\widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)\left(x_{I_1'},x_{I_2'}\right)\right|^2\left|I\left(x_{I_1'},x_{I_2'},x_{I_1},x_{I_2}\right)\right|\Big\}^{\frac{1}{2}}$$

$$\times\Big\{\sum_{k_1\leq k_1'}\sum_{k_2\leq k_2'}\sum_{I_1'}\sum_{I_2'}\sum_{I_1}\sum_{I_2}\mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2)$$

$$\cdot\left|\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1},x_{I_2}\right)\right|^2\left|I\left(x_{I_1'},x_{I_2'},x_{I_1},x_{I_2}\right)\right|\Big\}^{\frac{1}{2}}.$$

Note that by the estimates for $|I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})|$ in (2.11) we have

$$\sum_{I_1'} \sum_{I_2'} \mu_1(I_1')\mu_2(I_2')I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})| \le C2^{(k_1-k_1')\varepsilon_1}2^{(k_2-k_2')\varepsilon_2}$$

and similarly

$$\sum_{I_1} \sum_{I_2} \mu_1(I_1)\mu_2(I_2)|I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})| \le C2^{(k_1-k_1')\varepsilon_1}2^{(k_2-k_2')\varepsilon_2}.$$

Therefore,

$$\begin{aligned}
&\sum_{k_1\le k_1'} \sum_{k_2\le k_2'} \sum_{I_1'} \sum_{I_2'} \sum_{I_1} \sum_{I_2} \mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2) \\
&\quad \times \left|\widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)(x_{I_1'}, x_{I_2'})\right|^2 |I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})| \\
&\le C\sum_{k_1\le k_1'} \sum_{k_2\le k_2'} 2^{(k_1-k_1')\varepsilon_1}2^{(k_2-k_2')\varepsilon_2} \sum_{I_1'} \sum_{I_2'} \mu_1(I_1')\mu_2(I_2')\left|\widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)(x_{I_1'}, x_{I_2'})\right|^2 \\
&\le C\sum_{k_1'} \sum_{k_2'} \sum_{I_1'} \sum_{I_2'} \mu_1(I_1')\mu_2(I_2')\left|\widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)(x_{I_1'}, x_{I_2'})\right|^2.
\end{aligned}$$

The last series above, by a result established in [14], is dominated by the constant times $\|g\|_2^2$. Similarly,

$$\begin{aligned}
&\sum_{k_1\le k_1'} \sum_{k_2\le k_2'} \sum_{I_1'} \sum_{I_2'} \sum_{I_1} \sum_{I_2} \mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2) \\
&\quad \times \left|\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1}, x_{I_2})\right|^2 |I(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2})| \le C\|f\|_2^2.
\end{aligned}$$

We thus conclude that $|\langle g, Tf\rangle_{\text{Case 1.1}}| \le C\|f\|_2\|g\|_2$. The estimate for $|\langle g, Tf\rangle_{\text{Case 2.1}}|$ is the same.

2.2.2. *Carleson measure estimate on $\widetilde{M} = M_1 \times M_2$*

in this subsection we handle bilinear form $\langle g, Tf\rangle_{\text{Case1.4}}$. The estimate of this term will be achieved by applying the Carleson measure estimate on $\widetilde{M} = M_1 \times M_2$. To see this, we first write

$$\begin{aligned}
IV\left(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}\right) &= \int D^1_{k_1'}\left(x_{I_1'}, u_1\right) D^2_{k_2'}\left(x_{I_2'}, u_2\right) \\
&\quad \times K(u_1, u_2, v_1, v_2)\, D^1_{k_1}\left(x_{I_1'}, x_{I_1}\right) \\
&\quad \times D^2_{k_2}\left(x_{I_2'}, x_{I_2}\right) d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2) \\
&= D^1_{k_1'}D^2_{k_2'}(T1)\left(x_{I_1'}, x_{I_2'}\right) D^1_{k_1}\left(x_{I_1'}, x_{I_1}\right) D^2_{k_2}\left(x_{I_2'}, x_{I_2}\right).
\end{aligned}$$

Thus we rewrite $\langle g, Tf\rangle_{\text{Case 1.4}}$ by

$$\sum_{k_1'}\sum_{k_2'}\sum_{I_1'}\sum_{I_2'}\mu_1\left(I_1'\right)\mu_2\left(I_2'\right)\widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)\left(x_{I_1'},x_{I_2'}\right)$$

$$\times D^1_{k_1'}D^2_{k_2'}(T1)\left(x_{I_1'},x_{I_2'}\right)S_{k_1'}S_{k_2'}(f)\left(x_{I_1'},x_{I_2'}\right),$$

where for $x_1, y_1 \in M_1$, $S_{k_1'}(x_1,y_1) = \sum_{k_1\le k_1'}\sum_{I_1}\mu(I_1)D^1_{k_1}(x_1,x_{I_1})\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)$ and similarly for $S_{k_2'}(x_2,y_2)$ on M_2.

In order to apply the Carleson measure estimate to $\langle g, Tf\rangle_{\text{Case 1.4}}$, we claim that $S_{k_1'}(x_1,y_1)$, the kernel of $S_{k_1'}$, satisfies the following estimate

$$|S_{k_1'}(x_1,y_1)| \le C\frac{1}{V_{2^{-k_1'}}(x_1)+V_{2^{-k_1'}}(y_1)+V(x_1,y_1)}\left(\frac{2^{-k_1'}}{2^{-k_1'}+d_1(x_1,y_1)}\right)^{\vartheta'}.$$

Similarly, $S_{k_2'}(x_2,y_2)$, the kernel of $S_{k_2'}$, satisfies the same estimate above with interchanging k_1', k_2'; x_1, x_2 and y_1, y_2, respectively.

Assuming the claim for the moment, then applying the Cauchy–Schwartz inequality yields

$$|\langle g, Tf\rangle_{\text{Case 1.4}}|$$

$$\le\left\{\sum_{k_1'}\sum_{k_2'}\sum_{I_1'}\sum_{I_2'}\mu_1(I_1')\mu_2(I_2')\left|\widetilde{\widetilde{D}}^1{}_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2'}(g)\left(x_{I_1'},x_{I_2'}\right)\right|^2\right\}^{\frac12}$$

$$\times\left\{\sum_{k_1'}\sum_{k_2'}\sum_{I_1'}\sum_{I_2'}\mu_1(I_1')\mu_2(I_2')\left|D^1_{k_1'}D^2_{k_2'}(T1)\left(x_{I_1'},x_{I_2'}\right)\right|^2\left|S_{k_1'}S_{k_2'}(f)\left(x_{I_1'},x_{I_2'}\right)\right|^2\right\}^{\frac12}.$$

Thus the first series above, by the discrete Littlewood-Paley L^2 estimate, is bounded by a constant times $\|g\|_2$. And the second series is bounded by $C\|f\|_2$ by applying the Carleson measure estimate on $\widetilde{M}$ since $T1 \in BMO(\widetilde{M})$ and hence $\mu_1(I_1')\mu_2(I_2')|D^1_{k_1'}D^2_{k_2'}(T1)(x_1,x_2)|^2$ is a Carleson measure on $\widetilde{M}\times\{\mathbb{Z}\times\mathbb{Z}\}$.

We now show the claim. To do this, we first consider the case when $d_1(x_1, y_1) < 2^{-k_1'}$. Then

$$\Big|\sum_{k_1\le k_1',\ d_1(x_1,y_1)<2^{-k_1'}}\ \sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1)\Big|$$
$$\le C\sum_{k_1\le k_1',\ d_1(x_1,y_1)<2^{-k_1'}}\frac{1}{V_{2^{-k_1}}(x_1)+V_{2^{-k_1}}(y_1)+V(x_1,y_1)}\left(\frac{2^{-k_1}}{2^{-k_1}+d_1(x_1,y_1)}\right)^{\vartheta'}$$
$$\le C\frac{1}{V_{2^{-k_1'}}(x_1)+V_{2^{-k_1'}}(y_1)+V(x_1,y_1)}\left(\frac{2^{-k_1'}}{2^{-k_1'}+d_1(x_1,y_1)}\right)^{\vartheta'}, \tag{2.12}$$

where $(\vartheta_1)'$ is the order of $\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_1, y_1)$. Next, we consider the case when $d_1(x_1, y_1) \ge 2^{-k_1'}$. Note first that

$$\sum_{k_1\le k_1'}\sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(f)(x_{I_1})$$
$$+\sum_{k_1>k_1'}\sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(f)(x_{I_1}) = f(x_1)$$

for all $f \in L^2(M_1)$ and the series converge in the norm of L^2. This implies that

$$\sum_{k_1\le k_1'}\sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1)$$
$$+\sum_{k_1>k_1'}\sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1) = \delta(x_1, y_1), \tag{2.13}$$

where we use δ to denote the Dirac function. Consequently, when $d_1(x_1, y_1) \ge 2^{-k_1'}$,

$$\Big|\sum_{k_1\le k_1',d_1(x_1,y_1)\ge 2^{-k_1'}}\ \sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1)\Big|$$
$$=\Big|\sum_{k_1>k_1',d_1(x_1,y_1)\ge 2^{-k_1'}}\ \sum_{I_1}\mu_1(I_1)D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1)\Big|$$
$$\le C\frac{1}{V_{2^{-k_1'}}(x_1)+V_{2^{-k_1'}}(y_1)+V(x_1,y_1)}\left(\frac{2^{-k_1'}}{2^{-k_1'}+d_1(x_1,y_1)}\right)^{\vartheta'},$$

where the last inequality follows from similar estimates in (2.12) and hence the claim is proved.

2.2.3. *Almost orthogonal estimate on M_1 and Carleson measure estimate on M_2*

In this subsection we only estimate $\langle g,Tf\rangle_{\text{Case 1.2}}$ since all proofs for $\langle g,Tf\rangle_{\text{Case 1.3}}$, $\langle g,Tf\rangle_{\text{Case 2.2}}$ and $\langle g,Tf\rangle_{\text{Case 2.3}}$ are similar to the proof of $\langle g,Tf\rangle_{\text{Case 1.2}}$. We first write

$$
\begin{aligned}
&II(x_{I_1'},x_{I_2'},x_{I_1},x_{I_2})\\
&=\int D^1_{k_1'}(x_{I_1'},u_1)D^2_{k_2'}(x_{I_2'},u_2)K(u_1,u_2,v_1,v_2)\\
&\quad\times[D^2_{k_2}(v_2,x_{I_2})-D^2_{k_2}(x_{I_2'},x_{I_2})]d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2)D^1_{k_1}(x_{I_1'},x_{I_1})\\
&\quad+IV(x_{I_1'},x_{I_2'},x_{I_1},x_{I_2})\\
&=\langle D^2_{k_2'}(x_{I_2'},u_2),\langle D^1_{k_1'}(x_{I_1'},\cdot),K_2(u_2,v_2)(1)\rangle\\
&\quad\times[D^2_{k_2}(v_2,x_{I_2})-D^2_{k_2}(x_{I_2'},x_{I_2})]\rangle D^1_{k_1}(x_{I_1'},x_{I_1})\\
&\quad+IV(x_{I_1'},x_{I_2'},x_{I_1},x_{I_2}).
\end{aligned}
$$

Now we set

$$
\begin{aligned}
J_{k_2',k_2}(u_2,v_2)=\sum_{k_1'}\sum_{I_1'}&\mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}\big(\widetilde{\widetilde{D}}^2{}_{k_2'}(g)(\cdot,x_{I_2'})\big)(x_{I_1'})\\
&\times\langle D^1_{k_1'}(x_{I_1'},\cdot),K_2(u_2,v_2)(1)\rangle S_{k_1'}\big(\widetilde{\widetilde{D}}^2{}_{k_2}(f)(\cdot,x_{I_2})\big)(x_{I_1'}),
\end{aligned}
$$

where $S_{k_1'}$ is defined as in Subsection 2.3.2.

As in Subsection 2.3.2, summing up for k_1' and I_1' and using the notation $J_{k_2',k_2}(u_2,v_2)$, we can rewrite $\langle g,Tf\rangle_{\text{Case 1.2}}$ as

$$
\begin{aligned}
&\langle g,Tf\rangle_{\text{Case 1.2}}\\
&=\sum_{k_2\le k_2'}\sum_{I_2'}\sum_{I_2}\mu_2(I_2')\mu_2(I_2)d\mu^2(u_2)d\mu^2(v_2)\\
&\quad\times\int\widetilde{\widetilde{D}}^2{}_{k_2'}(x_{I_2'},u_2)J_{k_2',k_2}(u_2,v_2)[D^1_{k_2}(v_2,x_{I_2})-D^1_{k_2}(x_{I_2'},x_{I_2})]+\langle g,Tf\rangle_{\text{Case 1.4}}.
\end{aligned}
$$

Therefore, it suffices to estimate the above series since the estimate $|\langle g,Tf\rangle_{\text{Case 1.4}}|\le C\|f\|_2\|g\|_2$ has been proved in Subsection 2.3.1. To this end, we claim that for fixed k_2' and k_2, $J_{k_2',k_2}(u_2, v_2)$ is a Calderón-Zygmund kernel on M_2 and the corresponding operator has the weak boundedness property. Moreover,

$$|J_{k_2',k_2}(u_2, v_2)|_{CZ} \le C\|\widetilde{\widetilde{D}}^2{}_{k_2'}(g)(\cdot, x_{I_2'})\|_2\|\widetilde{\widetilde{D}}^2{}_{k_2}(f)(\cdot, x_{I_2})\|_2. \tag{2.14}$$

Assuming the claim for the moment, by the almost orthogonality argument as in Subsection 2.3.1 we obtain

$$\begin{aligned}
&\Bigg|\sum_{k_2\le k_2'}\sum_{I_2'}\sum_{I_2}\mu_2(I_2')\mu_2(I_2)\\
&\times\int \widetilde{\widetilde{D}}^2{}_{k_2'}\left(x_{I_2'},u_2\right)J_{k_2',k_2}(u_2,v_2)\left[D^2_{k_2}(v_2,x_{I_2})-D^2_{k_2}\left(x_{I_2'},x_{I_2}\right)\right]d\mu^2(u_2)d\mu^2(v_2)\Bigg|\\
&\le C\sum_{k_2\le k_2'}\sum_{I_2'}\sum_{I_2}\mu_2(I_2')\mu_2(I_2)|J_{k_2',k_2}(u_2, v_2)|_{CZ}\\
&\quad\times 2^{-(k_2-k_2')\varepsilon_2}\frac{1}{V_{2^{-k_2}}\left(x_{I_2'}\right)+V_{2^{-k_2}}\left(x_{I_2}\right)+V\left(x_{I_2'},x_{I_2}\right)}\frac{2^{-k_2\varepsilon_2}}{\left(2^{-k_2}+d_2\left(x_{I_2'},x_{I_2}\right)\right)^{\varepsilon_2}}
\end{aligned}$$

which, by a similar estimate as in Subsection 2.3.1, implies that the above series is dominated by a constant times

$$\begin{aligned}
&\sum_{k_2\le k_2'}\sum_{I_2'}\sum_{I_2}\mu_2(I_2')\mu_2(I_2)2^{-(k_2-k_2')\varepsilon_2}\frac{1}{V_{2^{-k_2}}\left(x_{I_2'}\right)+V_{2^{-k_2}}\left(x_{I_2}\right)+V\left(x_{I_2'},x_{I_2}\right)}\\
&\quad\times\frac{2^{-k_2\varepsilon_2}}{\left(2^{-k_2}+d_2\left(x_{I_2'},x_{I_2}\right)\right)^{\varepsilon_2}}\left\|\widetilde{\widetilde{D}}^2{}_{k_2'}(g)\left(\cdot,x_{I_2'}\right)\|_2\|\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(\cdot,x_{I_2}\right)\right\|_2\\
&\le C\|f\|_2\|g\|_2.
\end{aligned}$$

Now we prove the claim for $J_{k_2',k_2}(u_2, v_2)$. We first denote by J_{k_2',k_2} the operator on M_2 associated with the kernel $J_{k_2',k_2}(u_2, v_2)$. We verify that J_{k_2',k_2} satisfies the weak boundedness property. In fact, using the weak boundedness property of T on M_2, *i.e.* (1.7), and the one-parameter discrete Carleson measure estimate, we have

$$\left|\left\langle J_{k_2',k_2}\phi^2,\psi^2\right\rangle\right|\le CV_{r_2}\left(x_2^0\right)\left\|\widetilde{\widetilde{D}}^2{}_{k_2'}(g)\left(\cdot,x_{I_2'}\right)\right\|_{L^2(M_1)}\left\|\widetilde{\widetilde{D}}^2{}_{k_2'}(f)\left(\cdot,x_{I_2'}\right)\right\|_{L^2(M_1)}$$

for all $\phi^2, \psi^2 \in A_{M_2}(\delta, x_2^0, r_2)$, where the set $A_{M_2}(\delta, x_2^0, r_2)$ is defined in Section 1. Next we verify that $J_{k_2',k_2}(u_2, v_2)$ satisfies the size and smoothness properties as defined in Subsection 3.1. Using the one-parameter discrete Carleson measure estimate again we can obtain that

$$
\begin{aligned}
&|J_{k_2',k_2}(u_2, v_2)| \\
&\le C\|K_2(u_2, v_2)(1)\|_{BMO(M_1)}\|\widetilde{\widetilde{D^1}}_{k_2'}(g)(\cdot, x_{I_2'})\|_{L^2(M_1)}\|\widetilde{\widetilde{D^2}}_{k_2'}(f)(\cdot, x_{I_2'})\|_{L^2(M_1)} \\
&\le C\frac{1}{V(u_2, v_2)}\|\widetilde{\widetilde{D^2}}_{k_2'}(g)(\cdot, x_{I_2'})\|_{L^2(M_1)}\|\widetilde{\widetilde{D^2}}_{k_2'}(f)(\cdot, x_{I_2'})\|_{L^2(M_1)}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&|J_{k_2',k_2}(u_2, v_2) - h_{k_2',k_2}(u_2', v_2)| \\
&\le C\|K_2(u_2, v_2)(1) - K_2(u_2', v_2)(1)\|_{CZ}\|\widetilde{\widetilde{D^2}}_{k_2'}(g)(\cdot, x_{I_2'})\|_{L^2(M_1)} \\
&\quad \times \|\widetilde{\widetilde{D^2}}_{k_2'}(f)(\cdot, x_{I_2'})\|_{L^2(M_1)} \\
&\le C\Big(\frac{d_2(u_2, u_2')}{d_2(u_2, v_2)}\Big)^{\varepsilon_2}\frac{1}{V(u_2, v_2)}\|\widetilde{\widetilde{D^2}}_{k_2'}(g)(\cdot, x_{I_2'})\|_{L^2(M_1)}\|\widetilde{\widetilde{D^2}}_{k_2'}(f)(\cdot, x_{I_2'})\|_{L^2(M_1)}
\end{aligned}
$$

for $d_2(u_2, u_2') \le \frac{1}{2A_2}d_2(u_2, v_2)$. The same estimate holds with u_2 and v_2 interchanged. Combining the estimates above, we get that $J_{k_2',k_2}(u_2, v_2)$ is a Calderón-Zygmund kernel on M_2 and hence (2.14) holds. The claim is concluded.

2.2.4. *Littlewood-Paley estimate on M_1 and Carleson measure estimate on M_2*

In this subsection we deal with $\langle g, Tf\rangle_{\text{Case 2.4}}$. We first write

$$
\begin{aligned}
VIII(x_{I_1'}, x_{I_2'}, x_{I_1}, x_{I_2}) = -\int & D^1_{k_1'}(x_{I_1'}, u_1)D^2_{k_2'}(x_{I_2'}, x_{I_2}) \\
&\times K(u_1, u_2, v_1, v_2)D^1_{k_1}(x_{I_1'}, x_{I_1}) \\
&\times D^1_{k_2}(v_2, x_{I_2})d\mu^1(u_1)d\mu^2(u_2)d\mu^1(v_1)d\mu^2(v_2) \\
= -D^1_{k_1'}D^2_{k_2}&\big((\widetilde{T})^*1\big)(x_{I_1'}, x_{I_2})D^1_{k_1}(x_{I_1'}, x_{I_1})D^2_{k_2'}(x_{I_2'}, x_{I_2}).
\end{aligned}
$$

We would like to point out that the partial adjoint operator $\widetilde{T}$ appears and will play a crucial role in the estimate for $\langle g, Tf\rangle_{\text{Case 2.4}}$. This is why $\widetilde{T}$ and $\widetilde{T}^*$ have to be taken into account in the proof of the sufficient conditions of Theorem 1.

To estimate $\langle g, Tf\rangle_{\text{Case 2.4}}$ we rewrite

$$\begin{aligned}
&\langle g, Tf\rangle_{\text{Case 2.4}}\\
&= -\sum_{k_1\le k_1'}\sum_{k_2>k_2'}\sum_{I_1'}\sum_{I_2'}\sum_{I_1}\sum_{I_2}\mu_1(I_1')\mu_1(I_1)\mu_2(I_2')\mu_2(I_2)\\
&\quad\times D^2_{k_2'}\left(x_{I_2'}, x_{I_2}\right) D^1_{k_1}\left(x_{I_1'}, x_{I_1}\right)\widetilde{\widetilde{D}}^1{}_{k_1}\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1}, x_{I_2}\right)\\
&\quad\times D^1_{k_1'}D^2_{k_2}\left(\left(\widetilde{T}\right)^* 1\right)\left(x_{I_1'}, x_{I_2}\right)\\
&= -\sum_{k_1'}\sum_{k_2}\sum_{I_1'}\sum_{I_2}\mu_1\left(I_1'\right)\mu_2(I_2)\widetilde{\widetilde{D}}^1{}_{k_1'}S_{k_2}(g)\left(x_{I_1'}, x_{I_2}\right) S_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1'}, x_{I_2}\right)\\
&\quad\times D_{k_1'}D_{k_2}\left(\left(\widetilde{T}\right)^* 1\right)\left(x_{I_1'}, x_{I_2}\right),
\end{aligned}$$

where the operators $S_{k_1'}$ and S_{k_2} are defined as in Subsection 2.3.2.

In order to estimate the last series above, for a $BMO(\widetilde{M})$ function b we introduce an operator W_b by the bilinear form $\langle g, W_b f\rangle$ which equals

$$\begin{aligned}
&\sum_{k_1'}\sum_{k_2}\sum_{I_1'}\sum_{I_2}\mu_1(I_1')\mu_2(I_2)\widetilde{\widetilde{D}}^1{}_{k_1'}S_{k_2}(g)\left(x_{I_1'}, x_{I_2}\right) S_{k_1'}\\
&\quad\times\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1'}, x_{I_2}\right) D^1_{k_1'}D^2_{k_2}(b)\left(x_{I_1'}, x_{I_2}\right).
\end{aligned}$$

It is easy to see that when $b = (\widetilde{T})^*1 \in BMO(\widetilde{M})$ and $\langle g, W_b f\rangle = -\langle g, Tf\rangle_{\text{Case 2.4}}$. Thus, we only need to show that for each $b \in BMO(\widetilde{M})$ the operator W_b is bounded on L^2, which would imply that $|\langle g, Tf\rangle_{\text{Case 2.4}}| \le C\|f\|_2\|g\|_2$. For this purpose, following an idea in [17] and interchanging the positions of functions f and b we define the operator $V_f(b) = W_b(f)$ and will prove that for each fixed $f \in L^\infty$ the operator V_f is a singular integral operator and bounded on L^2. Moreover, there exists a constant C independent of f such that for all $b \in L^2$,

$$\|V_f(b)\|_2 \le C\|f\|_\infty\|b\|_2.$$

Furthermore, we will show that V_f satisfies the conditions in Theorem 3 and thus, V_f is also bounded on $BMO(\widetilde{M})$ satisfying

$$\|V_f(b)\|_{BMO} \le C\|f\|_\infty\|b\|_{BMO}.$$

We can rewrite the above estimate by

$$\|W_b(f)\|_{BMO} \le C\|f\|_\infty\|b\|_{BMO}$$

for each $b \in BMO(\widetilde{M})$ and all $f \in L^\infty(\widetilde{M})$.

This means that for each $b \in BMO(\widetilde{M})$ the operator W_b is a bounded operator from L^∞ to $BMO(\widetilde{M})$. Similarly, the operator W_b^*, the adjoint operator of W_b, is a bounded operator from L^∞ to $BMO(\widetilde{M})$ since W_b and W_b^* satisfy the same conditions. Finally, by the duality argument and interpolation, W_b is bounded on L^2 and hence, as mentioned, the bilinear form $\langle g, Tf\rangle_{\text{Case 2.4}}$ is bounded by a constant times $\|f\|_2\|g\|_2$.

To achieve this goal, we will show that for each fixed $f \in L^\infty$, V_f is a singular integral operator as defined in Section 1 and moreover, there exists a constant C independent of f and $b \in L^2$ such that

$$\|V_f(b)\|_2 \le C\|f\|_\infty\|b\|_2.$$

We first prove that V_f is bounded on L^2. To this end, for $g \in L^2$, we write

$$\begin{aligned}\langle g, V_f(b)\rangle = &\sum_{k_1'}\sum_{k_2}\sum_{I_1'}\sum_{I_2}\mu_1(I_1')\mu_2(I_2)\widetilde{\widetilde{D}}^1{}_{k_1'}S_{k_2}(g)\left(x_{I_1'}, x_{I_2}\right)S_{k_1'}\\ &\times \widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1'}, x_{I_2}\right)D^1_{k_1'}D^2_{k_2}(b)\left(x_{I_1'}, x_{I_2}\right).\end{aligned}$$

Note that if $f \in L^\infty$ then $S_{k_1'}(f)(x_{I_1'}, \cdot)$ is also a bounded function on M_2 for fixed k_1' and I_1' with

$$\|S_{k_1'}(f)\left(x_{I_1'}, \cdot\right)\|_\infty \le C\|f\|_\infty.$$

Thus $\mu_2(I_2)|\widetilde{\widetilde{D}}_{k_2}\big(S_{k_1'}(f)(x_{I_1'}, \cdot)\big)(x_{I_2})|^2$ is a Carleson measure on $M_2\times k_2$ uniformly for all k_1' and $x_{I_1'} \in M_1$. Therefore,

$$\begin{aligned}\big|\langle g, V_f(b)\rangle\big| = &\Bigg|\sum_{k_1'}\sum_{I_1'}\mu_1(I_1')\Bigg[\sum_{k_2}\sum_{\tau_2}\mu_2(I_2)S_{k_2}\left(\widetilde{\widetilde{D}}^1{}_{k_1}(g)\left(x_{I_1'}, \cdot\right)\right)\left(x_{I_2}\right)\\ &\times D^2_{k_2}\left(D^1_{k_1'}(b)\left(x_{I_1'}, \cdot\right)\right)\left(x_{I_2}\right)\widetilde{\widetilde{D}}^2{}_{k_2}\left(S_{k_1'}(f)\left(x_{I_1'}, \cdot\right)\right)\left(x_{I_2}\right)\Bigg]\Bigg|\\ \le &\sum_{k_1'}\sum_{I_1'}\mu_1\left(I_1'\right)\left\|\widetilde{\widetilde{D}}^1{}_{k_1}(g)\left(x_{I_1'}, \cdot\right)\right\|_{L^2(M_2)}\\ &\times\left\|D^1_{k_1'}(b)\left(x_{I_1'}, \cdot\right)\right\|_{L^2(M_2)}\left\|S_{k_1'}(f)\left(x_{I_1'}, \cdot\right)\right\|_{L^\infty(M_2)}\\ \le &\, C\|f\|_{L^\infty(\widetilde{M})}\left(\sum_{k_1'}\sum_{I_1'}\mu_1(I_1')\left\|\widetilde{\widetilde{D}}^1{}_{k_1}(g)\left(x_{I_1'}, \cdot\right)\right\|^2_{L^2(M_2)}\right)^{1/2}\\ &\times\left(\sum_{k_1'}\sum_{I_1'}\mu_1(I_1')\left\|D^1_{k_1'}(b)\left(x_{I_1'}, \cdot\right)\right\|^2_{L^2(M_2)}\right)^{1/2}\\ \le &\, C\|f\|_{L^\infty(\widetilde{M})}\|g\|_{L^2(\widetilde{M})}\|b\|_{L^2(\widetilde{M})},\end{aligned}$$

which, by taking the supremum for all $\|g\|_2 \le 1$, implies that V_f is bounded on $L^2(\widetilde{M})$ with $\|V_f\|_{L^2\to L^2} \le C\|f\|_{L^\infty}$.

To verify that V_f is a singular integral operator as defined in Section 1, we can consider V_f as a pair $\big((V_f)_1, (V_f)_2\big)$ of operators on M_2 and M_1, respectively, such that

$$\langle g_1 \otimes g_2, V_f h_1 \otimes h_2\rangle = \iint g_1(x_1)\langle g_2, (V_f)_1(x_1, y_1)h_2\rangle h_1(y_1)d\mu^1(x_1)d\mu^1(y_1)$$

for all $g_1, h_1 \in C_0^\eta(M_1)$ and $g_2, h_2 \in C_0^\eta(M_2)$ with $\operatorname{supp} g_1 \cap \operatorname{supp} h_1 = \varnothing$ and

$$\langle g_1 \otimes g_2, V_f h_1 \otimes h_2\rangle = \iint g_2(x_2)\langle g_1, (V_f)_2(x_2, y_2)h_1\rangle h_2(y_2)d\mu^2(x_2)d\mu^2(y_2)$$

for all $g_1, h_1 \in C_0^\eta(M_1)$ and $g_2, h_2 \in C_0^\eta(M_2)$ with $\operatorname{supp} g_2 \cap \operatorname{supp} h_2 = \varnothing$.

It suffices to show that $(V_f)_i(x_i, y_i)$, $i = 1, 2$, satisfies properties (i), (ii) and (iii) of singular integral operator given in Section 1. We need only to verify $(V_f)_1(x_1, y_1)$ since the estimates for $(V_f)_2(x_2, y_2)$ are similar.

Note that for any fixed x_1, y_1 on M_1, $(V_f)_1(x_1, y_1)$ is an operator on M_2 associated with the kernel $(V_f)_1(x_1, y_1)(x_2, y_2)$ which is equal to $V_f(x_1, x_2, y_1, y_2)$. We recall that

$$\|(V_f)_1(x_1, y_1)\|_{CZ} = \|(V_f)_1(x_1, y_1)\|_{L^2(M_2)\to L^2(M_2)} + |(V_f)_1(x_1, y_1)|_{CZ(M_2)},$$

where $|(V_f)_1(x_1, y_1)|_{CZ(M_2)}$ is the smallest constant such that the inequalities (a), (b) and (c) in Section 1 hold for the kernel $(V_f)_1(x_1, y_1)(x_2, y_2)$ when x_1, y_1 are fixed and $x_2, y_2 \in M_2$. Therefore, to verify that $(V_f)_1(x_1, y_1)$ satisfies properties (i), (ii) and (iii) of singular integral operator, all we need to do is to show the following estimates:

(I) $\|(V_f)_1(x_1, y_1)\|_{L^2\to L^2} \le C\|f\|_{L^\infty} \dfrac{1}{V(x_1, y_1)}$;

(II) $\|(V_f)_1(x_1, y_1) - (V_f)_1(x_1, y_1')\|_{L^2\to L^2}$

$$\le C\|f\|_{L^\infty} \left(\frac{d_1(y_1, y_1')}{d_1(x_1, y_1)}\right)^{\varepsilon_1} \frac{1}{V(x_1, y_1)}$$

if $d_1(y_1, y_1') \le d_1(x_1, y_1)/2A_1$.

Similarly for interchanging x_1 and y_1;

(III) $|(V_f)_1(x_1, y_1)(x_2, y_2)| \le C\|f\|_{L^\infty(\widetilde{M})} \dfrac{1}{V(x_1, y_1)} \dfrac{1}{V(x_2, y_2)}$;

(IV) $|(V_f)_1(x_1, y_1)(x_2, y_2) - (V_f)_1(x_1, y_1)(x_2', y_2)|$

$$\le C\|f\|_{L^\infty(\widetilde{M})} \frac{1}{V(x_1, y_1)} \left(\frac{d_2(x_2, x_2')}{d_2(x_2, y_2)}\right)^{\varepsilon_2} \frac{1}{V(x_2, y_2)}$$

if $d_2(x_2, x_2') \le d_2(x_2, y_2)/2A_2$.

Similarly for interchanging x_2 and y_2;

(V) $|(V_f)_1(x_1,y_1)(x_2,y_2)-(V_f)_1(x_1',y_1)(x_2,y_2)|$

$$\le C\|f\|_{L^\infty(\widetilde{M})}\left(\frac{d_1(x_1,x_1')}{d_1(x_1,y_1)}\right)^{\varepsilon_1}\frac{1}{V(x_1,y_1)}\frac{1}{V(x_2,y_2)}$$

if $d_1(y_1,y_1')\le d_1(x_1,y_1)/2A_1$.

Similarly for interchanging x_1 and y_1;

(VI) $\big|\big[(V_f)_1(x_1,y_1)(x_2,y_2)-(V_f)_1(x_1',y_1)(x_2,y_2)\big]$

$$-\big[(V_f)_1(x_1,y_1)(x_2',y_2)-(V_f)_1(x_1',y_1)(x_2',y_2)\big]\big|$$

$$\le C\|f\|_{L^\infty(\widetilde{M})}\left(\frac{d_1(x_1,x_1')}{d_1(x_1,y_1)}\right)^{\varepsilon_1}\frac{1}{V(x_1,y_1)}\left(\frac{d_2(x_2,x_2')}{d_2(x_2,y_2)}\right)^{\varepsilon_2}\frac{1}{V(x_2,y_2)}$$

if $d_1(x_1,x_1')\le d_1(x_1,y_1)/2A_1$ and $d_2(x_2,x_2')\le d_2(x_2,y_2)/2A_2$.

Similarly for interchanging x_2 and y_2, or interchanging x_1 and y_1.

To see (I), for fixed $x_1, y_1\in M_1$ we have

$$\|(V_f)_1(x_1,y_1)\|_{L^2\to L^2}=\sup_{g_2:\,\|g_2\|_{L^2(M_2)}\le 1}\ \sup_{h_2:\,\|h_2\|_{L^2(M_2)}\le 1}|\langle h_2,(V_f)_1(x_1,y_1)g_2\rangle|$$

$$=\sup_{g_2:\,\|g_2\|_{L^2(M_2)}\le 1}\ \sup_{h_2:\,\|h_2\|_{L^2(M_2)}\le 1}\Big|\sum_{k_1'}\sum_{I_1'}\mu_1(I_1')\widetilde{\widetilde{D}}^1{}_{k_1'}\left(x_1,x_{I_1'}\right)D^1_{k_1'}\left(x_{I_1'},y_1\right)$$

$$\times\Big[\sum_{k_2}\sum_{I_2}\mu_2(I_2)S_{k_2}(h_2)\left(x_{I_2}\right)D^2_{k_2}(g_2)\left(x_{I_2}\right)S_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1'},x_{I_2}\right)\Big]\Big|$$

$$\le C\|f\|_{L^\infty}\sup_{g_2:\,\|g_2\|_{L^2(M_2)}\le 1}\ \sup_{h_2:\,\|h_2\|_{L^2(M_2)}\le 1}\|h_2\|_{L^2(M_2)}\|g_2\|_{L^2(M_2)}$$

$$\times\sum_{k_1'}\sum_{I_1'}\mu_1\left(I_1'\right)\left|\widetilde{\widetilde{D}}^1{}_{k_1'}\left(x_1,x_{I_1'}\right)\right|\left|D^1_{k_1'}\left(x_{I_1'},y_1\right)\right|$$

$$\le C\|f\|_{L^\infty}\frac{1}{V(x_1,y_1)},$$

where in the first inequality we first apply Schwartz's inequality and then use the Littlewood-Paley estimate on L^2 for g_2 and the fact that if $f\in L^\infty$ then $\mu_2(I_2)\cdot|D^2_{k_2}(S_{k_1'}f)(x_{I_1'},x_{I_2})|^2$ is a Carleson measure on $M_2\times k_2$ uniformly for all k_1' and all $x_{I_1'}\in M_1$. Moreover, the Carleson measure norm of $\mu_2(I_2)|D^2_{k_2}(S_{k_1'}f)(x_{I_1'},x_{I_2})|^2$ is bounded by some constant times $\|f\|_{L^\infty}$. The last inequality follows from the standard estimate.

To verify (II), for $d_1(y_1, y_1') \le d_1(x_1, y_1)/2A_1$ and $\|g_2\|_{L^2(M_2)}, \|h_2\|_{L^2(M_2)} \le 1$,

$$
\begin{aligned}
&|\langle h_2, [(V_f)_1(x_1, y_1) - (V_f)_1(x_1, y_1')]g_2\rangle| \\
&\quad = \Big| \sum_{k_1'} \sum_{I_1'} \mu_1(I_1') \widetilde{\widetilde{D}}^1{}_{k_1'}(x_1, x_{I_1'})[D^1_{k_1'}(x_{I_1'}, y_1) - D^1_{k_1'}(x_{I_1'}, y_1')] \\
&\qquad \times \Big[\sum_{k_2} \sum_{I_2} \mu_2(I_2) S_{k_2}(h_2)(x_{I_2}) D^2_{k_2}(g_2)(x_{I_2}) S_{k_1'} \widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1'}, x_{I_2}) \Big] \Big|.
\end{aligned}
$$

Applying the smoothness property of $D^1_{k_1'}(x_{I_1'}, y_1)$ and the same proof above for the second series yields

$$
|\langle h_2, [(V_f)_1(x_1, y_1) - (V_f)_1(x_1, y_1')]g_2\rangle| \le C\|f\|_{L^\infty} \left(\frac{d_1(y_1, y_1')}{d_1(x_1, y_1)} \right)^{\varepsilon_1} \frac{1}{V(x_1, y_1)},
$$

which, by taking the supremum over all $\|g_2\|_{L^2(M_2)}, \|h_2\|_{L^2(M_2)} \le 1$ implies

$$
\begin{aligned}
&\|(V_f)_1(x_1, y_1) - (V_f)_1(x_1, y_1')\|_{L^2 \to L^2} \\
&\quad \le C\|f\|_{L^\infty} \left(\frac{d_1(y_1, y_1')}{d_1(x_1, y_1)} \right)^{\varepsilon_1} \frac{1}{V(x_1, y_1)}.
\end{aligned} \tag{2.15}
$$

Similarly, (2.15) holds with interchanging x_1 and y_1.

We now turn to estimate (III). This follows directly from the following standard estimate.

$$
\begin{aligned}
&|(V_f)_1(x_1, y_1)(x_2, y_2)| \\
&\quad \le \sum_{k_1'} \sum_{k_2} \sum_{I_1'} \sum_{I_2} \mu_1\left(I_1'\right) \mu_2(I_2) \left| \widetilde{\widetilde{D}}^1{}_{k_1'} \left(x_1, x_{I_1'}\right) S_{k_2}\left(x_2, x_{I_2}\right) \right| \\
&\qquad \times \left| S_{k_1'} \widetilde{\widetilde{D}}^2{}_{k_2}(f) \left(x_{I_1'}, x_{I_2}\right) \right| \left| D^1_{k_1'}(x_{I_1'}, y_1) D^1_{k_2}\left(x_{I_2}, y_2\right) \right| \\
&\quad \le \sum_{k_1'} \sum_{k_2} \sum_{I_1'} \sum_{I_2} \mu_1(I_1') \mu_2(I_2) \left| \widetilde{\widetilde{D}}^1{}_{k_1'} \left(x_1, x_{I_1'}\right) D^1_{k_1'} \left(x_{I_1'}, y_1\right) \right| \\
&\qquad \times \left| S_{k_2}\left(x_2, x_{I_2}\right) D^2_{k_2}\left(x_{I_2}, y_2\right) \right| \\
&\quad \le C\|f\|_{L^\infty(\widetilde{M})} \frac{1}{V(x_1, y_1)} \frac{1}{V(x_2, y_2)}.
\end{aligned} \tag{2.16}
$$

To estimate (IV), for $d_2(x_2, x_2') \le d_2(x_2, y_2)/2A_2$ we write

$$\begin{aligned}&\left|(V_f)_1(x_1, y_1)(x_2, y_2) - (V_f)_1(x_1, y_1)(x_2', y_2)\right|\\ &\quad\le \sum_{k_1'}\sum_{k_2}\sum_{I_1'}\sum_{I_2} \mu_1(I_1')\mu_2(I_2)\left|\widetilde{\widetilde{D}}^1{}_{k_1'}\left(x_1, x_{I_1'}\right)\left[S_{k_2}\left(x_2, x_{I_2}\right) - S_{k_2}\left(x_2', x_{I_2}\right)\right]\right|\\ &\qquad\times\left|S_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2}(f)\left(x_{I_1'}, x_{I_2}\right)\right|\left|D^1_{k_1'}\left(x_{I_1'}, y_1\right)D^2_{k_2}\left(x_{I_2}, y_2\right)\right|.\end{aligned}$$

We **claim** that $S_{k_2}(x_2, x_{I_2})$, which is defined in Subsection 2.3.2, satisfies the following smoothness estimate.

$$\begin{aligned}&|S_{k_2}(x_2, x_{I_2}) - S_{k_2}(x_2', x_{I_2})|\\ &\le C\left(\frac{d_2(x_2, x_2')}{2^{-k_2} + d_2(x_2, x_{I_2})}\right)^{\varepsilon_2}\frac{1}{V_{2^{-k_2}}(x_2) + V(x_2, x_{I_2})}\left(\frac{2^{-k_2}}{2^{-k_2} + d_2(x_2, x_{I_2})}\right)^{\varepsilon_2} \qquad (2.17)\end{aligned}$$

for $\varepsilon_2 < \vartheta_2$ and $d_2(x_2, x_2') < (2^{-k_2} + d_2(x_2, x_{I_2}))/2A_2$. We assume (2.17) first and then obtain

$$\begin{aligned}&\left|(V_f)_1(x_1, y_1)(x_2, y_2) - (V_f)_1(x_1, y_1)(x_2', y_2)\right|\\ &\quad\le C\|f\|_{L^\infty(\widetilde{M})}\frac{1}{V(x_1, y_1)}\left(\frac{d_2(x_2, x_2')}{d_2(x_2, y_2)}\right)^{\varepsilon_2}\frac{1}{V(x_2, y_2)}. \qquad (2.18)\end{aligned}$$

Similarly, (2.18) holds with interchanging x_2 and y_2. The estimates in (2.16) and (2.18) imply

$$|(V_f)_1(x_1, y_1)|_{CZ} \le C\|f\|_{L^\infty(\widetilde{M})}\frac{1}{V(x_1, y_1)}. \qquad (2.19)$$

Next, we turn to verify the estimate in (V). For $d_1(x_1, x_1') \le d_1(x_1, y_1)/2A_1$ We write

$$\begin{aligned}&(V_f)_1(x_1, y_1)(x_2, y_2) - (V_f)_1(x_1', y_1)(x_2, y_2)\\ &\quad= \sum_{k_1'}\sum_{k_2}\sum_{I_1'}\sum_{I_2}\mu_1(I_1')\mu_2(I_2)\left[\widetilde{\widetilde{D}}^1{}_{k_1'}(x_1, x_{I_1'}) - \widetilde{\widetilde{D}}^1{}_{k_1'}(x_1', x_{I_1'})\right]\\ &\qquad\times S_{k_2}(x_2, x_{I_2})S_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1'}, x_{I_2})D^1_{k_1'}(x_{I_1'}, y_1)D^2_{k_2}(x_{I_2}, y_2).\end{aligned}$$

As in the proof of (2.18), instead of using the smoothness estimate for $S_{k_2}(x_2, x_{I_2})$, applying the smoothness condition of $\widetilde{D^1}_{k_1'}$, we get

$$\begin{aligned}&\left|(V_f)_1(x_1,y_1)(x_2,y_2)-(V_f)_1(x_1',y_1)(x_2,y_2)\right|\\&\quad\le C\|f\|_{L^\infty(\widetilde{M})}\left(\frac{d_1(x_1,x_1')}{d_1(x_1,y_1)}\right)^{\varepsilon_1}\frac{1}{V(x_1,y_1)}\frac{1}{V(x_2,y_2)}.\end{aligned}\tag{2.20}$$

Similarly, (2.20) holds with interchanging x_1 and y_1.

Finally, to see (VI), for $d_2(x_2, x_2') \le d_2(x_2, y_2)/2A_2$ we have

$$\begin{aligned}&\Big|\big[(V_f)_1(x_1,y_1)(x_2,y_2)-(V_f)_1(x_1',y_1)(x_2,y_2)\big]\\&\quad-\big[(V_f)_1(x_1,y_1)(x_2',y_2)-(V_f)_1(x_1',y_1)(x_2',y_2)\big]\Big|\\&\quad=\Big|\sum_{k_1'}\sum_{k_2}\sum_{I_1'}\sum_{I_2}\mu_1(I_1')\mu_2(I_2)\Big[\widetilde{\widetilde{D^1}}_{k_1'}(x_1,x_{I_1'})-\widetilde{\widetilde{D^1}}_{k_1'}(x_1',x_{I_1'})\Big]\\&\qquad\times[S_{k_2}(x_2,x_{I_2})-S_{k_2}(x_2',x_{I_2})]\\&\qquad\times S_{k_1'}\widetilde{\widetilde{D^2}}_{k_2}(f)(x_{I_1'},x_{I_2})D^1_{k_1'}(x_{I_1'},y_1)D^2_{k_2}(x_{I_2},y_2)\Big|\\&\quad\le C\|f\|_{L^\infty(\widetilde{M})}\left(\frac{d_1(x_1,x_1')}{d_1(x_1,y_1)}\right)^{\varepsilon_1}\frac{1}{V(x_1,y_1)}\left(\frac{d_2(x_2,x_2')}{d_2(x_2,y_2)}\right)^{\varepsilon_2}\frac{1}{V(x_2,y_2)},\end{aligned}\tag{2.21}$$

where in the last inequality we use the smoothness property of $\widetilde{\widetilde{D^1}}_{k_1'}$ and (2.17). Similarly, (2.21) holds with interchanging x_2 and y_2 or x_1 and y_1.

All the estimates of (2.20) and (2.21) give

$$\begin{aligned}&\left|\big[(V_f)_1(x_1,y_1)(x_2,y_2)-(V_f)_1(x_1',y_1)(x_2,y_2)\big]\right|_{CZ}\\&\quad\le C\|f\|_{L^\infty(\widetilde{M})}\left(\frac{d_1(x_1,x_1')}{d_1(x_1,y_1)}\right)^{\varepsilon_1}\frac{1}{V(x_1,y_1)}.\end{aligned}\tag{2.22}$$

Similarly, (2.22) holds with interchanging x_1 and y_1.

As a consequence, (2.19) and (2.22) yield that $(V_f)_1(x_1, y_1)$ satisfies the properties (i), (ii) and (iii) given in Section 1 for the kernel K_1. It remains to show the claim, that is, the estimate in (2.17). Indeed, when $d_2(x_2, x_{I_2}) < 2^{-k_2}$ and

$d_1(x_2, x_2') < (2^{-k_2} + d_2(x_2, x_{I_2}))/2A_2$, we have

$$
\begin{aligned}
&|S_{k_2}(x_2, x_{I_2}) - S_{k_2}(x_2', x_{I_2})| \\
&\quad = \Big| \sum_{k_2' \le k_2,\ d_2(x_2, x_{I_2}) < 2^{-k_2}} \sum_{I_2'} \mu(I_2') D^2_{k_2'}(x_2, x_{I_2'}) \widetilde{\widetilde{D^2}}_{k_2'}(x_{I_2'}, x_{I_2}) \\
&\qquad - \sum_{k_2' \le k_2,\ d_2(x_2, x_{I_2}) < 2^{-k_1'}} \sum_{I_2'} \mu(I_2') D^2_{k_2'}(x_2', x_{I_2'}) \widetilde{\widetilde{D^1}}_{k_1}(x_{I_2'}, x_{I_2}) \Big| \\
&\quad \le C \sum_{k_2' \le k_2,\ d_2(x_2, x_{I_2}) < 2^{-k_2}} \left(\frac{d_2(x_2, x_2')}{2^{-k_2'} + d_2(x_2, x_{I_2})} \right)^{\varepsilon_2} \\
&\qquad \times \frac{1}{V_{2^{-k_2'}}(x_2) + V(x_2, x_{I_2})} \left(\frac{2^{-k_2'}}{2^{-k_2'} + d_2(x_2, x_{I_2})} \right)^{\varepsilon_2} \\
&\quad \le C \left(\frac{d_2(x_2, x_2')}{2^{-k_2} + d_2(x_2, x_{I_2})} \right)^{\varepsilon} \frac{1}{V_{2^{-k_2}}(x_2) + V(x_2, x_{I_2})} \left(\frac{2^{-k_2}}{2^{-k_2} + d_2(x_2, x_{I_2})} \right)^{\varepsilon}.
\end{aligned}
$$

Next, we consider the case when $d_2(x_2, x_{I_2}) \ge 2^{-k_2}$ and $d_2(x_2, x_2') < (2^{-k_2} + d_2(x_2, x_{I_2}))/2A_2$. In this case, using the identity (2.13), we obtain

$$
\begin{aligned}
&\Big| \sum_{k_2' \le k_2, d_2(x_2, x_{I_2}) \ge 2^{-k_2}} \sum_{I_2'} \mu(I_2') D^2_{k_2'}(x_2, x_{I_2'}) \widetilde{\widetilde{D^2}}_{k_2'}(x_{I_2'}, x_{I_2}) \\
&\qquad - \sum_{k_2' \le k_2, d_2(x_2, x_{I_2}) \ge 2^{-k_2}} \sum_{I_2'} \mu(I_2') D_{k_2'}(x_2', x_{I_2'}) \widetilde{\widetilde{D^2}}_{k_2'}(x_{I_2'}, x_{I_2}) \Big| \\
&\le \Big| \sum_{k_2' > k_2, d_2(x_2, x_{I_2}) \ge 2^{-k_2}} \sum_{I_2'} \mu(I_2') D^2_{k_2'}(x_2, x_{I_2'}) \widetilde{\widetilde{D^2}}_{k_2'}(h)(x_{I_2'}, x_{I_2}) \\
&\qquad - \sum_{k_2' > k_2, d_2(x_2, x_{I_2}) \ge 2^{-k_2}} \sum_{I_2'} \mu(I_2') D_{k_2'}(x_2', x_{I_2'}) \widetilde{\widetilde{D^2}}_{k_2'}(h)(x_{I_2'}, x_{I_2}) \Big| \\
&\le C \left(\frac{d_2(x_2, x_2')}{2^{-k_2} + d_2(x_2, x_{I_2})} \right)^{\varepsilon_2} \frac{1}{V_{2^{-k_2}}(x_2) + V(x_2, x_{I_2})} \left(\frac{2^{-k_2}}{2^{-k_2} + d_2(x_2, x_{I_2})} \right)^{\varepsilon_2},
\end{aligned}
$$

which implies the claim.

Now we have proved that V_f is a product Calderón-Zygmund operator with $\|V_f\|_{L^2 \to L^2} \le C\|f\|_{L^\infty}$. In order to apply Theorem 3 given in next section to show that V_f is bounded on $BMO(\widetilde{M})$, we only need to verify that $(V_f)_1(1) = (V_f)_2(1) = 0$. To do this, we would like to recall the definition of $T_1(1) = T_2(1) =$

0 and $(T^*)_1(1) = (T^*)_2(1) = 0$ as defined in Section 1. $T_1(1) = 0$ is equivalent to $\langle g_1, \langle g_2, T_2 f_2\rangle 1\rangle = 0$ for all $g_1 \in C^\eta_{00}(M_1)$ and $f_2, g_2 \in C^\eta_0(M_2)$, that is, for $g_1 \in C^\eta_{00}(M_1)$, $g_2 \in C^\eta_{00}(M_2)$ and almost everywhere $y_2 \in M_2$,

$$\iint g(x_1)g(x_2)K(x_1, x_2, y_1, y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1) = 0.$$

While $T_1{}^*(1) = 0$ means $\langle g_2, T_2 f_2\rangle^* 1 = 0$ in the same conditions; that is, for $g_1 \in C^\eta_{00}(M_1)$, $g_2 \in C^\eta_{00}(M_2)$ and almost everywhere $x_2 \in M_2$,

$$\iint g(y_1)g(y_2)K(x_1, x_2, y_1, y_2)d\mu^1(x_1)d\mu^1(y_1)d\mu^2(y_2) = 0.$$

To verify $(V_f)_1(1) = 0$, for $g_1 \in C^\eta_{00}(M_1)$, $g_2 \in C^\eta_{00}(M_2)$ and almost everywhere $y_2 \in M_2$ we have

$$\begin{aligned}&\iint g(x_1)g(x_2)V_f(x_1, x_2, y_1, y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1)\\ &= \iint g(x_1)g(x_2)\sum_{k_1'}\sum_{I_1'}\sum_{k_2}\sum_{I_2}\mu_1(I_1')\mu_2(I_2)\widetilde{\widetilde{D}}^1{}_{k_1'}(x_1, x_{I_1'})S_{k_2}(x_2, x_{I_2})\\ &\quad\times S_{k_1'}\widetilde{\widetilde{D}}^2{}_{k_2}(f)(x_{I_1'}, x_{I_2})D^1_{k_1'}(x_{I_1'}, y_1)D^2_{k_2}(x_{I_2}, y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1) = 0,\end{aligned}$$

where the last equality follows from the fact that $\int D^1_{k_1'}(x_{I_1'}, y_1)dy_1 = 0$. Similarly for $(V_f)_2(1) = 0$. As mentioned, we conclude that $|\langle g, Tf\rangle_{\text{Case2.4}}| \le C\|f\|_2\|g\|_2$.

The proof of the sufficient conditions for Theorem 1 is complete and hence the proof of Theorem 1 is concluded.

3. Proofs of Theorems 2 and 3

In this section we prove Theorems 2 and 3. We first prove the "if" part of Theorem 2 by applying Theorem 2.5 for the vector-valued product Calderón-Zygmund operators. The "if" part of Theorem 3 then follows from the "if" part of Theorem 2 by the duality argument. To show the converse, we will prove the "only if" part of Theorem 3 first and the "only if" part of Theorem 2 then follows from the duality argument directly.

3.1. "If" part of Theorem 2

To show the "if" part of Theorem 2, note first that $L^2(\widetilde{M}) \cap H^p(\widetilde{M})$ is dense in $H^p(\widetilde{M})$, see [14] for this result, and thus it suffices to prove that if T is the L^2 bounded product Calderón-Zygmund operator on $\widetilde{M}$ with a pair kernel (K_1, K_2)

satisfying the conditions (i) – (iii) and $(T^*)_1(1) = (T^*)_2(1) = 0$ then there exists a positive constant C independent of f such that

$$\|Tf\|_{H^p} \leq C\|f\|_{H^p}$$

for all $f \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$.

By a result in [14], this is equivalent to showing that

$$\|\widetilde{S}(Tf)\|_p \leq C\|f\|_{H^p}, \tag{3.1}$$

where $\widetilde{S}(Tf)$ is defined by

$$\widetilde{S}(f)(x_1, x_2) = \left\{ \sum_{\ell_1=-\infty}^{\infty} \sum_{\ell_2=-\infty}^{\infty} |D^1_{\ell_1} D^2_{\ell_2}(f)(x_1, x_2)|^2 \right\}^{1/2}. \tag{3.2}$$

The crucial idea for (3.1) is that by using the discrete Calderón reproducing formula for $f \in L^2(\widetilde{M})$, we can write the term $D^1_{\ell_1} D^2_{\ell_2}(Tf)(x_1, x_2)$ in (3.2) as

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \sum_{I_1} \sum_{I_2} \mu_1(I_1)\mu_2(I_2)(D^1_{\ell_1} D^2_{\ell_2} T D^1_{k_1}(\cdot, x_{I_1}) D^2_{k_2}(\cdot, x_{I_2}))(x_1, x_2)$$
$$\times \widetilde{\widetilde{D^1}}_{k_1} \widetilde{\widetilde{D^2}}_{k_2}(f)(x_{I_1}, x_{I_2}),$$

where the fact that T is bounded on $L^2(\widetilde{M})$ is used. This leads to considering the Hilbert space $\mathcal{H}$ defined by

$$\mathcal{H} = \left\{ \{h_{\ell_1,\ell_2}\}_{\ell_1,\ell_2 \in \mathbb{Z}} : \|h_{\ell_1,\ell_2}\|_{\mathcal{H}} := \left(\sum_{\ell_1=-\infty}^{\infty} \sum_{\ell_2=-\infty}^{\infty} |h_{\ell_1,\ell_2}|^2 \right)^{1/2} < \infty \right\}.$$

We then rewrite the operator $\widetilde{S} \circ T$ as the $\mathcal{H}$-valued operator $\mathcal{L}_{\ell_1,\ell_2}$, whose kernel is defined as

$$\begin{aligned} &\mathcal{L}_{\ell_1,\ell_2}(x_1, x_2, y_1, y_2) \\ &= \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \sum_{I_1} \sum_{I_2} \mu_1(I_1)\mu_2(I_2) D^1_{\ell_1} D^2_{\ell_2} T D^1_{k_1} D^2_{k_2}(x_1, x_2, x_{I_1}, x_{I_2}) \\ &\quad \times \widetilde{\widetilde{D^1}}_{k_1}(x_{I_1}, y_1) \widetilde{\widetilde{D^2}}_{k_2}(x_{I_2}, y_2). \end{aligned} \tag{3.3}$$

Therefore, the estimate in (3.1) is equivalent to

$$\|\mathcal{L}_{\ell_1,\ell_2}(f)\|_{L^p_{\mathcal{H}}} \leq C\|f\|_{H^p}, \tag{3.4}$$

Note first that $\mathcal{L}_{\ell_1,\ell_2}$ is bounded from L^2 to $L^2_{\mathcal{H}}$ since $\widetilde{S}$ and T are both bounded on L^2. The idea to show (3.4) is to apply Theorem 2.5 with vector-valued version. For this purpose, it suffices to verify the following conditions:

(I) $$\big\|(\mathcal{L}_{\ell_1,\ell_2})_1(x_1,y_1)-(\mathcal{L}_{\ell_1,\ell_2})_1(x_1,y_1')\big\|_{L^2(M_2)\to L^2_{\mathcal{H}}(M_2)} \le C\left(\frac{d_1(y_1,y_1')}{d_1(x_1,y_1)}\right)^{(\varepsilon_1)'}\frac{1}{V(x_1,y_1)} \quad \text{if } d_1(y_1,y_1')\le d_1(x_1,y_1)/2A_1;$$

and similar result holds for $(\mathcal{L}_{\ell_1,\ell_2})_2(x_2,y_2)$ for the variable y_2;

(II) $$\Big|\big[\mathcal{L}_{\ell_1,\ell_2}(x_1,y_1,x_2,y_2)-\mathcal{L}_{\ell_1,\ell_2}(x_1,y_1',x_2,y_2)\big] - \big[\mathcal{L}_{\ell_1,\ell_2}(x_1,y_1,x_2,y_2')-\mathcal{L}_{\ell_1,\ell_2}(x_1,y_1',x_2,y_2')\big]\Big|_{\mathcal{H}} \le C\left(\frac{d_1(y_1,y_1')}{d_1(x_1,y_1)}\right)^{(\varepsilon_1)'}\frac{1}{V(x_1,y_1)}\left(\frac{d_2(y_2,y_2')}{d_2(x_2,y_2)}\right)^{(\varepsilon_2)'}\frac{1}{V(x_2,y_2)}$$

if $d_1(y_1,y_1')\le d_1(x_1,y_1)/2A_1$ and $d_2(y_2,y_2')\le d_2(x_2,y_2)/2A_2$.

To show (I), note that

$$\begin{aligned}
&\big\|(\mathcal{L}_{\ell_1,\ell_2})_1(x_1,y_1)-(\mathcal{L}_{\ell_1,\ell_2})_1(x_1,y_1')\big\|_{L^2(M_2)\to L^2_{\mathcal{H}}(M_2)}\\
&= \sup_{f:\,\|f\|_{L^2(M_2)}\le 1}\bigg(\int_{M_2}\Big\|\int_{M_2}\big[\mathcal{L}_{\ell_1,\ell_2}(x_1,x_2,y_1,y_2) - \mathcal{L}_{\ell_1,\ell_2}(x_1,x_2,y_1',y_2)\big]f(y_2)dy_2\Big\|_{\mathcal{H}}^2dx_2\bigg)^{1/2}.
\end{aligned}$$

We write

$$\begin{aligned}
&\bigg(\int_{M_2}\Big\|\int_{M_2}[\mathcal{L}_{\ell_1,\ell_2}(x_1,x_2,y_1,y_2)-\mathcal{L}_{\ell_1,\ell_2}(x_1,x_2,y_1',y_2)]f(y_2)dy_2\Big\|_{\mathcal{H}}^2dx_2\bigg)^{1/2}\\
&=\bigg(\int_{M_2}\sum_{\ell_1=-\infty}^{\infty}\sum_{\ell_2=-\infty}^{\infty}\bigg|D^2_{\ell_2}\bigg(\int\sum_{k_1=-\infty}^{\infty}\sum_{I_1}\mu_1(I_1)D^1_{\ell_1}(x_1,u_1)K(u_1,\cdot,v_1,v_2)\\
&\quad\times D^1_{k_1}(v_1,x_{I_1})\big[\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1},y_1)-\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1},y_1')\big]\\
&\quad\times f(v_2)d\mu^1(u_1)d\mu^1(v_1)d\mu^2(v_2)\bigg)(x_2)\bigg|^2d\mu^2(x_2)\bigg)^{1/2}\\
&\le C\bigg(\sum_{\ell_1=-\infty}^{\infty}\int_{M_2}\bigg|\int\sum_{k_1=-\infty}^{\infty}\sum_{I_1}\mu_1(I_1)D^1_{\ell_1}(x_1,u_1)K(u_1,x_2,v_1,v_2)\\
&\quad\times D^1_{k_1}(v_1,x_{I_1})\big[\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1},y_1)-\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1},y_1')\big]\\
&\quad\times f(v_2)d\mu^1(u_1)d\mu^1(v_1)d\mu^2(v_2)\bigg|^2d\mu^2(x_2)\bigg)^{1/2},
\end{aligned}\tag{3.5}$$

where the last inequality follows from the Littlewood-Paley estimate on M_2.

We claim that for any fixed ℓ_1 and $(\varepsilon_1)'$ with $(\varepsilon_1)' < \varepsilon_1$ there exists positive a constant C such that for $d_1(y_1, y_1') \le d_1(x_1, y_1)/2A_1$ and $\|f\|_2 \le 1$,

$$\begin{aligned}&\bigg(\int_{M_2}\bigg|\int \sum_{k_1=-\infty}^{\infty}\sum_{I_1}\mu_1(I_1)D^1_{\ell_1}(x_1,u_1)K(u_1,x_2,v_1,v_2)D^1_{k_1}(v_1,x_{I_1})\\&\quad\times\big[\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)-\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1')\big]f(v_2)d\mu^1(u_1)d\mu^1(v_1)d\mu^2(v_2)\bigg|^2 d\mu^2(x_2)\bigg)^{1/2}\\&\le C\left(\frac{d_1(y_1,y_1')}{2^{-\ell_1}}\right)^{(\varepsilon_1)'}\frac{1}{V_{2^{-\ell_1}}(x_1)+V(x_1,y_1)}\left(\frac{2^{-\ell_1}}{2^{-\ell_1}+d_1(x_1,y_1)}\right)^{(\varepsilon_1)'}.\end{aligned}\tag{3.6}$$

Assuming (3.6) and inserting (3.6) into (3.5) together with the following standard estimate

$$\begin{aligned}&\sum_{\ell_1}\left(\frac{d_1(y_1,y_1')}{2^{-\ell_1}}\right)^{2(\varepsilon_1)'}\left(\frac{1}{V_{2^{-\ell_1}}(x_1)+V(x_1,y_1)}\right)^2\left(\frac{2^{-\ell_1}}{2^{-\ell_1}+d_1(x_1,y_1)}\right)^{2(\varepsilon_1)'}\\&\quad\le C\Big(\frac{d_1(y_1,y_1')}{d_1(x_1,y_1)}\Big)^{2(\varepsilon_1)'}\frac{1}{V^2(x_1,y_1)}\end{aligned}$$

yields that for $d_1(y_1, y_1') \le d_1(x_1, y_1)/2A_1$ and $\|f\|_2 \le 1$,

$$\begin{aligned}&\bigg(\int_{M_2}\bigg\|\int_{M_2}[\mathcal{L}_{k_1',k_2'}(x_1,x_2,y_1,y_2)-\mathcal{L}_{k_1',k_2'}(x_1,x_2,y_1',y_2)]f(y_2)dy_2\bigg\|^2_{\mathcal{H}}dx_2\bigg)^{1/2}\\&\quad\le C\left(\frac{d_1(y_1,y_1')}{d_1(x_1,y_1)}\right)^{(\varepsilon_1)'}\frac{1}{V(x_1,y_1)},\end{aligned}$$

which implies (I).

In order to show the estimate in (3.6), we will apply the almost orthogonal argument. For this purpose, we write the left-hand side of (3.6) by $\sup_{h:\ \|h\|_{L^2(M_2)}\le 1}|H|$, where

$$\begin{aligned}H=&\int\sum_{k_1=-\infty}^{\infty}\sum_{I_1}\mu_1(I_1)D^1_{\ell_1}(x_1,u_1)\langle h,K_1(u_1,v_1)f\rangle D^1_{k_1}(v_1,x_{I_1})du_1dv_1\\&\times\big[\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)-\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1')\big].\end{aligned}$$

As in Subsection 2.3.1, for fixed ℓ_1 we decompose the summation over k_1 by $k_1 > \ell_1$ and $k_1 \le \ell_1$ and denote them by E and F, respectively.

Note that, as in Subsection 2.3.1, for $k_1 > \ell_1$, $\|f\|_2 \le 1$ and $\|g\|_2 \le 1$, the condition that $(T)_1^*(1) = 0$ implies the following almost orthogonal estimate

$$\Big| \int D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, v_1) f\rangle D^1_{k_1}(v_1, x_{I_1}) du_1 dv_1 \Big|$$

$$\le C 2^{-(k_1-\ell_1)(\varepsilon_1)'} \frac{1}{V_{2^{-\ell_1}}(x_1) + V_{2^{-\ell_1}}(x_{I_1}) + V(x_1, x_{I_1})} \frac{2^{-\ell_1(\varepsilon_1)'}}{(2^{-\ell_1} + d_1(x_1, x_{I_1}))^{(\varepsilon_1)'}},$$

which, together with the smoothness property of $\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1)$, yields that $|E|$ is bounded by

$$C \sum_{k_1 > \ell_1} \sum_{I_1} \mu_1(I_1) 2^{-(k_1-\ell_1)(\varepsilon_1)'} \frac{1}{V_{2^{-\ell_1}}(x_1) + V_{2^{-\ell_1}}(x_{I_1}) + V(x_{I_1}, x_1)}$$
$$\times \left(\frac{2^{-\ell_1}}{2^{-\ell_1} + d_1(x_1, x_{I_1})}\right)^{(\varepsilon_1)'} \left(\frac{d_1(y_1, y_1')}{2^{-k_1} + d_1(x_{I_1}, y_1)}\right)^{(\varepsilon_1)'}$$
$$\times \frac{1}{V_{2^{-k_1}}(x_{I_1}) + V(x_{I_1}, y_1)} \left(\frac{2^{-k_1}}{2^{-k_1} + d_1(x_{I_1}, y_1)}\right)^{(\varepsilon_1)'},$$

which gives the right-hand side of (3.6).

Similarly, we decompose F as

$$\begin{aligned} F = &\int \sum_{k_1 \le \ell_1} \sum_{I_1} \mu_1(I_1) D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, v_1) f\rangle \\ &\times [D^1_{k_1}(v_1, x_{I_1}) - D^1_{k_1}(x_1, x_{I_1})] d\mu^1(u_1) d\mu^1(v_1) \\ &\times \big[\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1) - \widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1')\big] \\ &+ \int \sum_{k_1 \le \ell_1} \sum_{I_1} \mu_1(I_1) D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, v_1) f\rangle \\ &\times D^1_{k_1}(x_1, x_{I_1}) d\mu^1(u_1) d\mu^1(v_1) \\ &\times \big[\widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1) - \widetilde{\widetilde{D}}{}^1{}_{k_1}(x_{I_1}, y_1')\big] \\ = &F_1 + F_2. \end{aligned}$$

Note that when $k_1 \le \ell_1$ we have the following almost orthogonal estimate that for $\|f\|_2 \le 1$ and $\|g\|_2 \le 1$,

$$\Big| \int D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, v_1) f\rangle [D^1_{k_1}(v_1, x_{I_1}) - D^1_{k_1}(x_1, x_{I_1})] d\mu^1(u_1) d\mu^1(v_1) \Big|$$

$$\le C 2^{-(\ell_1-k_1)(\varepsilon_1)'} \frac{1}{V_{2^{-k_1}}(x_1) + V_{2^{-k_1}}(x_{I_1}) + V(x_1, x_{I_1})} \frac{2^{-k_1(\varepsilon_1)'}}{(2^{-k_1} + d_1(x_1, x_{I_1}))^{(\varepsilon_1)'}}.$$

Therefore, F_1 satisfies the same estimate as E. To estimate F_2, we rewrite it as

$$F_2 = \sum_{k_1\le \ell_1}\sum_{I_1} \mu_1(I_1) D^1_{k_1}(x_1, x_{I_1})\big[\widetilde{\widetilde{D^1}}_{k_1}(x_{I_1}, y_1) - \widetilde{\widetilde{D^1}}_{k_1}(x_{I_1}, y_1')\big]$$
$$\times \int D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, \cdot) f\rangle(1) d\mu^1(u_1)$$
$$= \big[S_{\ell_1}(x_1, y_1) - S_{\ell_1}(x_1, y_1')\big] \int D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, \cdot) f\rangle(1) d\mu^1(u_1),$$

where for $x_1, y_1 \in M_1$, $S_{\ell_1}(x_1, y_1) = \sum_{k_1\le\ell_1}\sum_{I_1} \mu_1(I_1) D^1_{k_1}(x_1, x_{I_1})\widetilde{\widetilde{D^1}}_{k_1}(x_{I_1}, y_1)$ and similarly for $S_{\ell_1}(x_1, y_1')$. Note that $S_{\ell_1}(x_1, y_1)$ and $S_{\ell_1}(x_1, y_1')$ satisfy the size and smoothness properties as proved in Subsections 2.3.2 and 3.3.4, respectively. Similar to the argument in Subsection 3.3.3, $\langle h, K_1(u_1, \cdot) f\rangle(1)$, as a function of u_1, lies in $BMO(M_1)$ with $\|\langle h, K_1(u_1, \cdot) f\rangle(1)\|_{BMO(M_1)} \le C\|f\|_{L^2(M_2)}\|h\|_{L^2(M_2)}$. Hence $\left|\int D^1_{\ell_1}(x_1, u_1)\langle h, K_1(u_1, \cdot) f\rangle(1) d\mu^1(u_1)\right| \le C\|f\|_{L^2(M_2)}\|h\|_{L^2(M_2)}$, where the constant C is independent of ℓ_1 and x_1 since for any ℓ_1 and x_1, $D^1_{\ell_1}(x_1, u_1)$ is in $H^1(M_1)$ with $\|D_{\ell_1}(x_1, \cdot)\|_{H^1(M_1)}$ uniformly bounded. As a consequence, we have

$$|F_2| \le C\big|S_{\ell_1}(x_1, y_1) - S_{\ell_1}(x_1, y_1')\big|\|f\|_{L^2(M_2)}\|h\|_{L^2(M_2)}.$$

Thus, applying the size properties of $S_{\ell_1}(x_1, y_1)$ and $S_{\ell_1}(x_1, y_1')$ for the case $\ell_1 : 2^{-\ell_1} \le 2A_1 d_1(y_1, y_1')$ and the smoothness properties of $S_{\ell_1}(x_1, y_1)$ for the case $\ell_1 : 2^{-\ell_1} > 2A_1 d_1(y_1, y_1')$, we conclude that F_2 satisfies the same estimate as F_1 and hence (3.6) holds.

To verify (II), it suffices to show that there exist positive constants C, ε and ε' with $\varepsilon' < \varepsilon$, such that

$$\text{(II}'\text{)}\quad |\mathcal{L}_{\ell_1,\ell_2}(x_1, x_2, y_1, y_2) - \mathcal{L}_{\ell_1,\ell_2}(x_1, x_2, y_1', y_2) - \mathcal{L}_{\ell_1,\ell_2}(x_1, x_2, y_1, y_2') + \mathcal{L}_{\ell_1,\ell_2}(x_1, x_2, y_1', y_2')|$$
$$\le C\Big(\frac{d_1(y_1, y_1')}{2^{-\ell_1} + d_1(x_1, y_1)}\Big)^{(\varepsilon_1)'} \frac{1}{V_{2^{-\ell_1}}(x_1) + V_{2^{-\ell_1}}(y_1) + V(x_1, y_1)}$$
$$\times \frac{2^{-\ell_1(\varepsilon_1)'}}{(2^{-\ell_1} + d_1(x_1, y_1))^{(\varepsilon_1)'}}\Big(\frac{d_2(y_2, y_2')}{2^{-\ell_2} + d_2(x_2, y_2)}\Big)^{(\varepsilon_2)'}$$
$$\times \frac{1}{V_{2^{-\ell_2}}(x_2) + V_{2^{-\ell_2}}(y_2) + V(x_2, y_2)} \frac{2^{-\ell_2(\varepsilon_2)'}}{(2^{-\ell_2} + d_2(x_2, y_2))^{(\varepsilon_2)'}}$$

for $d_1(y_1, y_1') \le \frac{1}{2A_1}(2^{-\ell_1} + d_1(x_1, y_1))$ and $d_2(y_2, y_2') \le \frac{1}{2A_2}(2^{-\ell_2} + d_2(x_2, y_2))$.

Note that from (3.3), we can write the left-hand side of (II') as

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\sum_{I_1}\sum_{I_2}\mu_1(I_1)\mu_2(I_2)D^1_{\ell_1}D^2_{\ell_2}TD^1_{k_1}D^2_{k_2}(x_1,x_2,x_{I_1},x_{I_2})$$
$$\times\big[\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)-\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1')\big]\big[\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2},y_2)-\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2},y_2')\big]=:\mathbb{L}_{\ell_1,\ell_2}.$$

Then, to estimate $\mathbb{L}_{\ell_1,\ell_2}$, it suffices to estimate the term $D^1_{\ell_1}D^2_{\ell_2}TD^1_{k_1}D^2_{k_2}(x_1,x_2,x_{I_1},x_{I_2})$, which is exactly the same as what we have done in Subsection 3.3. To be more precise, for any fixed integers ℓ_1 and ℓ_2 we consider the following four cases: $\ell_1\geq k_1$ and $\ell_2\geq k_2$; $\ell_1\geq k_1$ and $\ell_2<k_2$; $\ell_1<k_1$ and $\ell_2\geq k_2$; $\ell_1<k_1$ and $\ell_2<k_2$. Then we write $\mathbb{L}_{\ell_1,\ell_2}=\mathbb{L}^1_{\ell_1,\ell_2}+\mathbb{L}^2_{\ell_1,\ell_2}+\mathbb{L}^3_{\ell_1,\ell_2}+\mathbb{L}^4_{\ell_1,\ell_2}$, where each $\mathbb{L}^i_{\ell_1,\ell_2}$ corresponds to each case. We now only consider $\mathbb{L}^1_{\ell_1,\ell_2}$ and $\mathbb{L}^2_{\ell_1,\ell_2}$ since the other two terms follow symmetrically.

For $\mathbb{L}^1_{\ell_1,\ell_2}$, following the Case 1 in Subsection 2.3, we decompose

$$D^1_{k_1'}D^2_{k_2'}TD^1_{k_1}D^2_{k_2}(x_1,x_2,x_{I_1},x_{I_2})$$
$$=: I(x_1,x_2,x_{I_1},x_{I_2})+II(x_1,x_2,x_{I_1},x_{I_2})$$
$$+III(x_1,x_2,x_{I_1},x_{I_2})+IV(x_1,x_2,x_{I_1},x_{I_2})$$

and then write $\mathbb{L}^1_{\ell_1,\ell_2}=\mathbb{L}^{1.1}_{\ell_1,\ell_2}+\mathbb{L}^{1.2}_{\ell_1,\ell_2}+\mathbb{L}^{1.3}_{\ell_1,\ell_2}+\mathbb{L}^{1.4}_{\ell_1,\ell_2}$, where

$$\mathbb{L}^{1.1}_{\ell_1,\ell_2}=\sum_{k_1\leq\ell_1}\sum_{k_2\leq\ell_2}\sum_{I_1}\sum_{I_2}\mu_1(I_1)\mu_2(I_2)I(x_1,x_2,x_{I_1},x_{I_2})$$
$$\times\Big[\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)-\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1')\Big]\Big[\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2},y_2)-\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2},y_2')\Big]$$

and similar for the other three terms.

For $\mathbb{L}^2_{\ell_1,\ell_2}$, since $(T^*)_2(1)=0$, the Case 2 in Subsection 2.3 gives

$$D^1_{k_1'}D^2_{k_2'}TD^1_{k_1}D^2_{k_2}(x_1,x_2,x_{I_1},x_{I_2})=V(x_1,x_2,x_{I_1},x_{I_2})+VI(x_1,x_2,x_{I_1},x_{I_2})$$

and hence we can write $\mathbb{L}^2_{\ell_1,\ell_2}=\mathbb{L}^{2.1}_{\ell_1,\ell_2}+\mathbb{L}^{2.2}_{\ell_1,\ell_2}$ similarly.

For $\mathbb{L}^{1.1}_{\ell_1,\ell_2}$ and $\mathbb{L}^{2.1}_{\ell_1,\ell_2}$, applying the almost orthogonality estimates for $I(x_1,x_2,x_{I_1},x_{I_2})$ and $V(x_1,x_2,x_{I_1},x_{I_2})$ as in Subsection 2.3.1, the smoothness properties for $\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)$ and $\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2},y_2)$, and then following the same proof as in Case 1.1 in Subsection 2.3.1, we conclude that $\mathbb{L}^{1.1}_{\ell_1,\ell_2}$ and $\mathbb{L}^{2.1}_{\ell_1,\ell_2}$ satisfies the estimate in (II').

For $\mathbb{L}^{1.4}_{\ell_1,\ell_2}$, applying the Carleson measure estimate for $IV(x_1,x_2,x_{I_1},x_{I_2})$ as in Case 1.4 in Subsection 2.3.2 and the smoothness properties for $\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1},y_1)$ and $\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2},y_2)$ implies that $\mathbb{L}^{1.4}_{\ell_1,\ell_2}$ is bounded by the right-hand side of (II').

Similarly, the almost orthogonality estimates on M_1 and the Carleson measure estimates on M_2 as in Case 1.2 in Subsection 2.3.3 and the smoothness properties of $\widetilde{\widetilde{D}}^1{}_{k_1}(x_{I_1}, y_1)$ and $\widetilde{\widetilde{D}}^2{}_{k_2}(x_{I_2}, y_2)$ gives the estimate in (II') for $\mathbb{L}^{1,2}_{\ell_1,\ell_2}$. Similarly for the estimate of $\mathbb{L}^{1,3}_{\ell_1,\ell_2}$. For more details of the proof, we refer the readers to the long version of this paper [13]. This finishes the proof of the "if" part of Theorem 2.

3.2. "If" part of Theorem 3

Note that if $f \in CMO^p(\widetilde{M})$, in general, $T(f)$ may not be well defined because f is a distribution in $\big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)\big)'$. The same problem appears in the proof of Theorem 2.5. The key fact used in the proof of Theorem 2.5 is that $L^2(\widetilde{M}) \cap H^p(\widetilde{M})$ is dense in $H^p(\widetilde{M})$. It turns out that to establish the boundedness of T on $H^p(\widetilde{M})$, it suffices to show the H^p boundedness of T for $f \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$. This method does not work for the present proof of the "If" part of Theorem 3 because $L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$ is not dense in $CMO^p(\widetilde{M})$. However, as a substitution, we have the following:

Lemma 3.1. *For* $\max\big(\frac{2\omega_1}{2\omega_1+\vartheta_1}, \frac{2\omega_2}{2\omega_2+\vartheta_2}\big) < p \le 1$, $L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$ *is dense in* $CMO^p(\widetilde{M})$ *in the weak topology* $(H^p(\widetilde{M}), CMO^p(\widetilde{M}))$. *More precisely, for each* $f \in CMO^p(\widetilde{M})$, *there exists a sequence* $\{f_n\} \subset L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$ *such that* $\|f_n\|_{CMO^p(\widetilde{M})} \le C\|f\|_{CMO^p(\widetilde{M})}$, *where* C *is a positive constant independent of* n *and* f, *and moreover, for each* $g \in H^p(\widetilde{M})$, $\langle f_n, g\rangle \to \langle f, g\rangle$ *as* $n \to \infty$.

See [19] for the proof.

We now prove the "if" part of Theorem 3. We first define T on $CMO^p(\widetilde{M})$ as follows. Given $f \in CMO^p(\widetilde{M})$, by Lemma 3.1, there is a sequence $\{f_n\} \subset L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$ such that $\|f_n\|_{CMO^p(\widetilde{M})} \le C\|f\|_{CMO^p(\widetilde{M})}$, and for each $g \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$, $\langle f_n, g\rangle \to \langle f, g\rangle$ as $n \to \infty$. Thus, for $f \in CMO^p(\widetilde{M})$, we define

$$\langle T(f), g\rangle := \lim_{n\to\infty} \langle T(f_n), g\rangle$$

for each $g \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$.

To see that this limit exists, we note that $\langle T(f_j - f_k), g\rangle = \langle f_j - f_k, T^*(g)\rangle$ since both $f_j - f_k$ and g belong to L^2 and T is bounded on L^2. T^* is bounded on L^2 and the kernel of T^* satisfies the conditions in Theorem 2. Moreover, $((T^*)_1)^*(1) = T_1(1) = 0$ and $((T^*)_2)^*(1) = T_2(1) = 0$. Therefore, by the "if" part of Theorem 2 which has been proved in Subsection 3.1, $T^*(g) \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$. Thus, by Lemma 3.1, $\langle f_j - f_k, T^*(g)\rangle$ tends to zero as $j, k \to \infty$. It is also easy to see that this limit is independent of the choice of the sequence f_n that satisfies the conditions in Lemma 3.1.

To finish the proof of "if" part of Theorem 3, we claim that for each $f \in L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$,

$$\|T(f)\|_{CMO^p(\widetilde{M})} \leq C\|f\|_{CMO^p(\widetilde{M})},$$

where the constant C is independent of f.

To see the above claim implies the "if" part of Theorem 3, by the definition of T on $CMO^p(\widetilde{M})$, for each $g \in L^2(\widetilde{M}) \cap H^p(\widetilde{M})$, $\langle T(f), g\rangle = \lim_{n\to\infty}\langle T(f_n), g\rangle$, where f_n satisfies the conditions in Lemma 3.1. Particularly, taking $g(x, y) = D^2_{k_2} D^1_{k_1}(x, y) \in \overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ and applying the claim yield

$$\begin{aligned}\|T(f)\|_{CMO^p(\widetilde{M})} &= \|\lim_{n\to\infty} T(f_n)\|_{CMO^p(\widetilde{M})} \leq \liminf_{n\to\infty} \|T(f_n)\|_{CMO^p(\widetilde{M})}\\ &\leq C\|f_n\|_{CMO^p(\widetilde{M})} \leq C\|f\|_{CMO^p(\widetilde{M})}.\end{aligned}$$

Thus, it remains to show the claim. The proof of the claim follows from the duality between $H^p(\widetilde{M})$ and $CMO^p(\widetilde{M})$, and the "if" part of Theorem 2. To be more precise, let $f \in L^2 \cap CMO^p(\widetilde{M})$ and $g \in L^2 \cap H^p(\widetilde{M})$. By the duality first and then the "if" part of Theorem 2, we have

$$\begin{aligned}|\langle T(f), g\rangle| &= |\langle f, T^*(g)\rangle|\\ &\leq \|f\|_{CMO^p(\widetilde{M})}\|T^*(g)\|_{H^p(\widetilde{M})} \leq C\|f\|_{CMO^p(\widetilde{M})}\|g\|_{H^p(\widetilde{M})}.\end{aligned}$$

This implies that for each $f \in L^2(\widetilde{M}) \cap CMO^p(\widetilde{M})$, $\ell_f(g) = \langle T(f), g\rangle$ defines a continuous linear functional on $L^2(\widetilde{M}) \cap H^p(\widetilde{M})$. Note that $L^2(\widetilde{M}) \cap H^p(\widetilde{M})$ is dense in $H^p(\widetilde{M})$. Thus, $\ell_f(g) = \langle T(f), g\rangle$ belongs to the dual of $H^p(\widetilde{M})$ and the norm of this linear functional is dominated by $C\|f\|_{CMO^p}$. By the duality of $H^p(\widetilde{M})$ with $CMO^p(\widetilde{M})$, again, there exists $h \in CMO^p(\widetilde{M})$ such that $\langle T(f), g\rangle = \langle h, g\rangle$ for each $g \in \overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2; \gamma_1, \gamma_2)$ and $\|h\|_{CMO^p} \leq C\|\ell_f\| \leq C\|f\|_{CMO^p(\widetilde{M})}$. The crucial fact we will use is that, taking $g(x, y) = D^2_{k_2} D^1_{k_1}(x, y)$, we obtain that $\langle T(f), D^2_{k_2} D^1_{k_1}\rangle = \langle h, D^2_{k_2} D^1_{k_1}\rangle$. Therefore, by the definition of $CMO^p(\widetilde{M})$, we have

$$\begin{aligned}&\|T(f)\|_{CMO^p(\widetilde{M})}\\ &= \sup_\Omega \left\{\frac{1}{|\Omega|^{\frac{2}{p}-1}} \sum_{k_1,k_2\in\mathbb{Z}} \sum_{I_1,I_2:I_1\times I_2\subset\Omega} |D^2_{k_2} D^1_{k_1}(T(f))(x_{I_1}, x_{I_2})|^2|I_1||I_2|\right\}^{1/2}\\ &= \sup_\Omega \left\{\frac{1}{|\Omega|^{\frac{2}{p}-1}} \sum_{k_1,k_2\in\mathbb{Z}} \sum_{I_1,I_2:I_1\times I_2\subset\Omega} |D^2_{k_2} D^1_{k_1}(h)(x_{I_1}, x_{I_2})|^2|I_1||I_2|\right\}^{1/2}\\ &= \|h\|_{CMO^p(\widetilde{M})} \leq C\|f\|_{CMO^p(\widetilde{M})}.\end{aligned}$$

The proof of the claim is concluded and hence the proof of "if " part of Theorem 3 is complete.

3.3. "Only if" parts of Theorems 2 and 3

We first show the "only if" part of Theorem 3. Suppose that T is a Calderón-Zygmund operator defined in Section 1 and bounded on $CMO^p(\widetilde{M})$. For each $f_2(x_2) \in C_0^\eta(M_2)$, we define the function $f(x_1,x_2)$ on $\widetilde{M}$ by $f(x_1,x_2) := \chi_1(x_1)f_2(x_2)$, where $\chi_1(x_1)=1$ on M_1. It is clear that f is in $CMO^p(\widetilde{M})$ with $\|f\|_{CMO^p(\widetilde{M})}=0$. Consequently, we have $Tf \in CMO^p(\widetilde{M})$ and $\|Tf\|_{CMO^p(\widetilde{M})} = 0$. Therefore,

$$\int_{M_2}\int_{M_1}\int_{M_2}\int_{M_1} g_1(x_1)g_2(x_2)K(x_1,y_1,x_2,y_2) \times f_2(y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1)d\mu^2(y_2)=0$$

for all $g_1\in C_0^\eta(M_1)$ with $\int g_1(x_1)d\mu^1(x_1)=0$, $g_2\in C_0^\eta(M_2)$ with $\int g_2(x_2)d\mu^2(x_2)=0$ and all $f_2\in C_0^\eta(M_2)$. Note that the above equality is equivalent to

$$\int_{M_2}\int_{M_1} T^*(g_1\otimes g_2)(y_1,y_2)f_2(y_2)d\mu^1(y_1)d\mu^2(y_2)=0.$$

Since T is bounded on $L^2(\widetilde{M})$, T^* is also bounded on $L^2(\widetilde{M})$. Therefore, $T^*(g_1\otimes g_2)\in L^1(\widetilde{M})\cap L^2(\widetilde{M})$ since $(g_1\otimes g_2)\in H^1(\widetilde{M})$. Note that $C_0^\eta(M_2)$ is dense in $L^2(M_2)$. This implies

$$\begin{aligned}\int_{M_1} &T^*(g_1\otimes g_2)(y_1,y_2)dy_1=0\\ &=\int_{M_1}\int_{M_2}\int_{M_1} g_1(x_1)g_2(x_2)K(x_1,y_1,x_2,y_2)d\mu^1(x_1)d\mu^2(x_2)d\mu^1(y_1)\end{aligned}$$

for all $g_1\in C_0^\eta(M_1)$ with $\int g_1(x_1)d\mu^1(x_1)=0$, $g_2\in C_0^\eta(M_2)$ with $\int g_2(x_2)d\mu^2(x_2)=0$ and for $y_2\in M_2$ almost everywhere. Thus, $T_1(1)=0$. Similarly we can prove that $T_2(1)=0$.

We now prove the "only if" part of Theorem 2. We claim that if T is bounded on L^2 and $H^p(\widetilde{M})$, then the adjoint operator T^* extends to a bounded operator from $CMO^p(\widetilde{M})$ to itself, where T^* is defined originally by $\langle Tf,g\rangle = \langle f,T^*g\rangle$ for all $f,g\in L^2(\widetilde{M})$.

To see this, let $f\in L^2(\widetilde{M})\cap H^p(\widetilde{M})$ and $g\in L^2(\widetilde{M})\cap CMO^p(\widetilde{M})$, then, by the duality between $H^p(\widetilde{M})$ and $CMO^p(\widetilde{M})$,

$$|\langle T^*g,f\rangle| = |\langle g,Tf\rangle| \le C\|f\|_{H^p(\widetilde{M})}\|g\|_{CMO^p(\widetilde{M})}.$$

This implies that $\langle T^*g,f\rangle$ defines a continuous linear functional on $H^p(\widetilde{M})$ because $L^2(\widetilde{M})\cap H^p(\widetilde{M})$ is dense in $H^p(\widetilde{M})$. Moreover, applying the same proof given in Subsection 4.2 yields

$$\|T^*g\|_{CMO^p(\widetilde{M})} \le C\|g\|_{CMO^p(\widetilde{M})}.$$

Then, applying the "only if" part of Theorem 3 for the operator T^* implies that $(T^*)_1(1)=(T^*)_2(1)=0$.

4. The case of n factors

In this section we first consider the case of 3 factors; that is, $\widetilde{M} = M_1 \times M_2 \times M_3$. We recall the definition of the Littlewood-Paley square function on $\widetilde{M}$.

Definition 4.1. Let $\{S^i_{k_i}\}_{k_i\in\mathbb{Z}}$ be approximations to the identity on M_i and $D^i_{k_i} = S^i_{k_i} - S^i_{k_i-1}$, $i = 1, 2, 3$. For $f \in \big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2, \beta_3; \gamma_1, \gamma_2, \gamma_3)\big)'$ with $0 < \beta_i, \gamma_i < \vartheta_i$, $i = 1, 2, 3$, $\widetilde{S}_d(f)$, the discrete Littlewood-Paley square function of f, is defined by

$$
\begin{aligned}
&S(f)(x_1, x_2, x_3)\\
&= \Big\{ \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \sum_{k_3=-\infty}^{\infty} \sum_{I_1^1} \sum_{I_2^2} \sum_{I_3^3} |D^1_{k_1} D^2_{k_2} D^3_{k_3}(f)(x_1, x_2, x_3)|^2 \\
&\qquad \times \chi_{I_1^1}(x_1)\chi_{I_2^2}(x_2)\chi_{I_3^3}(x_3)\Big\}^{1/2},
\end{aligned}
$$

where for each k_i, I^i_i ranges over all the dyadic cubes in M_i with side-length $\ell(I^i_i) = 2^{-k_i-N_i}$, and N_i are large fixed positive integers, for $i = 1, 2, 3$.

We recall the Hardy spaces H^p and Carleson measure spaces CMO^p on $\widetilde{M}$ as follows.

Definition 4.2 ([14]). Let $\max\big(\frac{\omega_1}{\omega_1+\vartheta_1}, \frac{\omega_2}{\omega_2+\vartheta_2}, \frac{\omega_3}{\omega_3+\vartheta_3}\big) < p \le 1$ and $0 < \beta_i, \gamma_i < \vartheta_i$ for $i = 1, 2, 3$.

$$
H^p(\widetilde{M}) := \big\{ f \in \big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2, \beta_3; \gamma_1, \gamma_2, \gamma_3)\big)' : \ S(f) \in L^p(\widetilde{M}) \big\}
$$

and if $f \in H^p(\widetilde{M})$, the norm (or quasi-norm) of f is defined by $\|f\|_{H^p(\widetilde{M})} = \|S(f)\|_p$.

Definition 4.3 ([14]). Let $\max\big(\frac{2\omega_1}{2\omega_1+\vartheta_1}, \frac{2\omega_2}{2\omega_2+\vartheta_2}, \frac{2\omega_3}{2\omega_3+\vartheta_3}\big) < p \le 1$ and $0 < \beta_i$, $\gamma_i < \vartheta_i$ for $i = 1, 2, 3$. Let $\{S^i_{k_i}\}_{k_i\in\mathbb{Z}}$ be approximations to the identity on M_i and for $k_i \in \mathbb{Z}$, set $D^i_{k_i} = S^i_{k_i} - S^i_{k_i-1}$, $i = 1, 2, 3$. The generalized Carleson measure space $CMO^p(\widetilde{M})$ is defined, for $f \in \big(\overset{\circ}{G}_{\vartheta_1,\vartheta_2}(\beta_1, \beta_2, \beta_3; \gamma_1, \gamma_2, \gamma_3)\big)'$, by

$$
\begin{aligned}
\|f\|_{CMO^p(\widetilde{M})} = \sup_{\Omega} \Big\{ &\frac{1}{\mu(\Omega)^{\frac{2}{p}-1}} \int_{\Omega} \sum_{k_1,k_2,k_3} \sum_{I_1^1\times I_2^2\times I_3^3 \subseteq \Omega} \big|D^1_{k_1} D^2_{k_2} D^3_{k_3}(f)(x_1, x_2, x_3)\big|^2 \\
&\times \chi_{I_1^1}(x_1)\chi_{I_2^2}(x_2)\chi_{I_3^3}(x_3)\, d\mu^1(x_1)\, d\mu^2(x_2)\, d\mu^3(x_3)\Big\}^{\frac{1}{2}} < \infty,
\end{aligned}
$$

where Ω are taken over all open sets in $\widetilde{M}$ with finite measures and for each k_i, I^i_i ranges over all the dyadic cubes in M_i with length $\ell(I^i_i) = 2^{-k_i-N_i}$, $i = 1, 2, 3$.

To consider singular integral operators on $\widetilde{M}$, we first introduce the space $C_0^\eta(\widetilde{M})$ by induction. Note that we have introduced $C_0^\eta(M_1 \times M_2)$ in Section 1. A function $f(x_1, x_2, x_3)$ is said to be in $C_0^\eta(\widetilde{M})$ if f has compact support and

$$\|f(x_1, x_2, \cdot)\|_{C_0^\eta(M_1\times M_2)} \in C_0^\eta(M_3).$$

Now we introduce a class of *product singular integral operators* on $\widetilde{M}$.

Let $T : C_0^\eta(\widetilde{M}) \to \big(C_0^\eta(\widetilde{M})\big)'$ be a linear operator with an associated distribution kernel $K(x_1, y_1, x_2, y_2, x_3, y_3)$, which is a continuous function on $\widetilde{M} \setminus \{(x_1, y_1, x_2, y_2, x_3, y_3) : x_i = y_i, \text{ for some } i, 1 \le i \le 3\}$. Moreover,

(i) $$\begin{aligned}&\langle T(\varphi_1 \otimes \varphi_2 \otimes \varphi_3), \psi_1 \otimes \psi_2 \otimes \psi_3\rangle \\ &\quad= \int K(x_1, y_1, x_2, y_2, x_3, y_3) \\ &\qquad\times \prod_1^3 \varphi_i(x_i)\psi_i(y_i)d\mu^1(x_1)d\mu^1(y_1)d\mu^2(x_2)d\mu^2(y_2)d\mu^3(x_3)d\mu^3(y_3)\end{aligned}$$

whenever φ_i and ψ_i are in $C_0^\eta(M_i)$ with disjoint supports, for $1 \le i \le 3$.

(ii) There exists a Calderón-Zygmund valued operator $K_3(x_3, y_3)$ on $M_1 \times M_2$ such that

$$\begin{aligned}&\langle T(\varphi_1 \otimes \varphi_2 \otimes \varphi_3), \psi_1 \otimes \psi_2 \otimes \psi_3\rangle \\ &\quad= \int \langle K_3(x_3, y_3)(\varphi_1 \otimes \varphi_2), \psi_1 \otimes \psi_2\rangle \varphi_3(x_3)\psi_3(y_3)d\mu^3(x_3)d\mu^3(y_3)\end{aligned}$$

whenever φ_i and ψ_i are in $C_0^\eta(M_i)$ for $1 \le i \le 3$ and $\operatorname{supp}\varphi_3 \cap \operatorname{supp}\psi_3 = \varnothing$. Moreover, $\|K_3(x_3, y_3)\|_{CZ(M_1\times M_2)}$ as a function of $x_3, y_3 \in M_3$, satisfies the following conditions:

(ii-a) $\|K_3(x_3, y_3)\|_{CZ,1,2} \le C\dfrac{1}{V(x_3, y_3)}$;

(ii-b) $\|K_3(x_3, y_3) - K_3(x_3, y_3')\|_{CZ,1,2} \le C\Big(\frac{d_3(y_3,y_3')}{d_3(x_3,y_3)}\Big)^{\varepsilon_3} \frac{1}{V(x_3,y_3)}$ if $d_3(y_3, y_3') \le \frac{d_3(x_3,y_3)}{2A_3}$;

(ii-c) $\|K_3(x_3, y_3) - K_3(x_3', y_3)\|_{CZ,1,2} \le C\Big(\dfrac{d_3(x_3, x_3')}{d_3(x_3, y_3)}\Big)^{\varepsilon_3} \dfrac{1}{V(x_3, y_3)}$ if $d_3(x_3, x_3') \le \frac{d_3(x_3,y_3)}{2A_3}$.

Here we use $\|\cdot\|_{CZ(M_1\times M_2)}$ to denote the Calderón-Zygmund norm of the product Calderón-Zygmund operators on $M_1 \times M_2$. More precisely, $\|T\|_{CZ(M_1\times M_2)} = \|T\|_{L^2\to L^2} + |K|_{CZ(M_1\times M_2)}$, where $|K|_{CZ,1,2} = \min(|K_1|_{CZ}, |K_2|_{CZ})$ by considering K as a pair (K_1, K_2) as in Section 1.

(iii) There exists a Calderón-Zygmund valued operator $K_{1,2}(x_1, y_1, x_2, y_2)$ on M_3 such that

$$\langle T(\varphi_1 \otimes \varphi_2 \otimes \varphi_3), \psi_1 \otimes \psi_2 \otimes \psi_3\rangle = \int \langle K_{1,2}(x_1, y_1, x_2, y_2)(\varphi_3), \psi_3\rangle \times \prod_{i=1}^{2} \varphi_i(x_i)\psi_i(y_i)d\mu^1(x_1)d\mu^1(y_1)d\mu^2(x_2)d\mu^2(y_2)$$

whenever φ_i and ψ_i are in $C_0^\eta(M_i)$ for $1 \le i \le 3$, and φ_i and ψ_i have disjoint supports for $i = 1, 2$. Moreover, as a function of (x_1, y_1, x_2, y_2), $K_{1,2}(x_1, y_1, x_2, y_2)$ satisfies the following conditions:

(iii-a) $\|K_{1,2}(x_1, y_1, x_2, y_2)\|_{CZ} \le C\dfrac{1}{V(x_1, y_1)}\dfrac{1}{V(x_2, y_2)}$;

(iii-b) $\|K_{1,2}(x_1, y_1, x_2, y_2) - K_{1,2}(x_1', y_1, x_2, y_2)\|_{CZ}$
$$\le C\Big(\frac{d_1(x_1, x_1')}{d_1(x_1, y_1)}\Big)^{\varepsilon_1}\frac{1}{V(x_1, y_1)}\frac{1}{V(x_2, y_2)} \quad \text{if } d_1(x_1, x_1') \le \frac{d_1(x_1, y_1)}{2A_1};$$

(iii-c) the condition (iii-b) also holds for interchanging x_1, x_2 with y_1, y_2;

(iii-d)
$$\|K_{1,2}(x_1, y_1, x_2, y_2) - K_{1,2}(x_1', y_1, x_2, y_2) - K_{1,2}(x_1, y_1, x_2', y_2) + K_{1,2}(x_1', y_1, x_2', y_2)\|_{CZ}$$
$$\le C\Big(\frac{d_1(x_1, x_1')}{d_1(x_1, y_1)}\Big)^{\varepsilon_1}\frac{1}{V(x_1, y_1)}\Big(\frac{d_2(x_2, x_2')}{d_2(x_2, y_2)}\Big)^{\varepsilon_2}\frac{1}{V(x_2, y_2)}$$
if $d_1(x_1, x_1') \le \frac{d_1(x_1,y_1)}{2A_1}$ and $d_2(x_2, x_2') \le \frac{d_2(x_2,y_2)}{2A_2}$

(iii-e) the condition (iii-d) also holds for interchanging x_1, x_2 with y_1, y_2.

(iv) The same conditions (ii) and (iii) hold for any permutation of the indices $1, 2, 3$. That is, we can consider T as a pair of $(K_{1,3}, K_2)$, as well as a pair of $(K_1, K_{2,3})$. Both K_1 and K_2 satisfy (ii). Similarly, both $K_{1,3}$ and $K_{2,3}$ satisfy (iii).

To state the result on $\widetilde{M}$, we need to deal with the partial adjoint operators $\widetilde{T}$. We have the following two classes of partial adjoint operators. For the first class, $\widetilde{T}_1$, the partial adjoint operator of T, is defined as

$$\langle \widetilde{T}_1(\varphi_1 \otimes \varphi_2 \otimes \varphi_3), \psi_1 \otimes \psi_2 \otimes \psi_3\rangle = \langle T(\psi_1 \otimes \varphi_2 \otimes \varphi_3), \varphi_1 \otimes \psi_2 \otimes \psi_3\rangle,$$

and similarly for $\widetilde{T}_2$ and $\widetilde{T}_3$. For the second class, $\widetilde{T}_{1,2}$, the partial adjoint operator of T, is defined as

$$\langle \widetilde{T}_{1,2}(\varphi_1 \otimes \varphi_2 \otimes \varphi_3), \psi_1 \otimes \psi_2 \otimes \psi_3\rangle = \langle T(\psi_1 \otimes \psi_2 \otimes \varphi_3), \varphi_1 \otimes \varphi_2 \otimes \psi_3\rangle,$$

and similarly $\widetilde{T}_{1,2}$ and $\widetilde{T}_{2,3}$. Thus, there are totally 6 partial adjoint operators.

We also define the weak boundedness property. Let T be a product singular integral operator on $\widetilde{M}$. We say that T has the WBP if

$$\|\langle K_1(\varphi_2 \otimes \varphi_3), \psi_2 \otimes \psi_3\rangle\|_{CZ(M_1)} \lesssim V_{r_2}(x_2^0)V_{r_3}(x_3^0)$$

$$\text{for all } \varphi_2, \psi_2 \in A_{M_2}(\delta, x_2^0, r_2),\ \varphi_3, \psi_3 \in A_{M_3}(\delta, x_3^0, r_3)$$

and

$$\|\langle K_{1,2}(\varphi_3), \psi_3\rangle\|_{CZ(M_1\times M_2)} \lesssim V_{r_3}(x_3^0) \qquad \text{for all } \varphi_3, \psi_3 \in A_{M_3}(\delta, x_3^0, r_3),$$

and the same conditions hold for K_1, K_2 and $K_{1,3}, K_{2,3}$, respectively.

Now we can state the result on $\widetilde{M} = M_1 \times M_2 \times M_3$.

Theorem 1′**.** Let T be a product singular integral operator on $\widetilde{M}$. Then T is bounded on $L^2(\widetilde{M})$ if and only if $T1, T^*1, \widetilde{T}_1 1, \widetilde{T}_2 1, \widetilde{T}_3 1, \widetilde{T}_{1,2} 1, \widetilde{T}_{1,3} 1$ and $\widetilde{T}_{2,3} 1$ lie in $BMO(\widetilde{M})$ and T has the weak boundedness property.

The general case $\widetilde{M} = M_1 \times \cdots \times M_n$ of n factors will follow by induction.

References

[1] A. CALDERÓN, *Intermediate spaces and interpolation, the complex method*, Studia Math. **24** (1964), 113–190.

[2] S.-Y. A. CHANG and R. FEFFERMAN, *A continuous version of the duality of H^1 and BMO on the bidisc*, Ann. of Math. **112** (1980), 179–201.

[3] S.-Y. A. CHANG and R. FEFFERMAN, *The Calderón-Zygmund decomposition on product domains*, Amer. J. Math. **104** (1982), 455–468.

[4] M. CHRIST, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), 601–628.

[5] D. DENG and Y. HAN, "Harmonic Analysis on Spaces of Homogeneous Type", Lecture Notes in Mathematics, Vol. 1966, Springer, Berlin, 2009.

[6] G. DAVID and J. L. JOURNÉ, *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math. **120** (1984), 371–397.

[7] G. DAVID, J. L. JOURNÉ and S. SEMMES, *Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation*, Rev. Mat. Iberoam. **1** (1985), 1–56.

[8] R. FEFFERMAN, *Bounded mean oscillation on the polydisk*, Ann. of Math. **110** (1979), 395-406.

[9] R. FEFFERMAN, *Harmonic analysis on product spaces*, Ann. of Math. **126** (1987), 109–130.

[10] R. FEFFERMAN and E. M. STEIN, *Singular integrals on product spaces*, Adv. Math. **45** (1982), 117–143.

[11] R. GUNDY and E. M. STEIN, *H^p theory for the polydisc*, Proc. Nat. Acad. Sci. **76** (1979), 1026–1029.

[12] Y. HAN, M.-Y. LEE, C.-C. LIN and Y.-C. LIN, *Calderón-Zygmund operators on product Hardy spaces*, J. Funct. Anal. **258** (2010), 2834-2861.

[13] Y. HAN, J. LI and C.-C. LIN, *$T1$ theorem on product Carnot-Caratheodory spaces*, ARXIV: 1209.6236.

[14] Y. HAN, J. LI and G. LU, *Multiparameter Hardy space theory on Carnot-Caratheodory spaces and product spaces of homogeneous type*, Trans. Amer. Math. Soc. **365** (2013), 319–360.
[15] Y. HAN, D. MÜLLER and D. YANG, *Littlewood-Paley-Stein characterizations for Hardy spaces on spaces of homogeneous type*, Math. Nachr. **279** (2006), 1505–1537.
[16] B. JESSEN, J. MARCINKIEWICZ and A. ZYGMUND, *Note on the differentiability of multiple integrals*, Funda. Math. **25** (1935), 217–234.
[17] J. L. JOURNÉ, *Calderón-Zygmund operators on product space*, Rev. Mat. Iberoam. **1** (1985), 55–92.
[18] J. L. JOURNÉ, *Two problems of Calderón-Zygmund theory on product spaces*, Ann. Inst. Fourier (Grenoble) **38** (1988), 111–132.
[19] J. LI and L. WARD, *Singular integrals on Carleson measure spaces CMO^p on product spaces of homogeneous type*, Proc. Amer. Math. Soc. **141** (2013), 2767–2782.
[20] Y. MEYER, "Wavelets and Operators", Cambridge University Press, 1992.
[21] A. NAGEL and E. M. STEIN, *The $\Box_b$-heat equation on pseudoconvex manifolds of finite type in $\mathbb{C}^2$*, Math. Z. **238** (2001), 37–88.
[22] A. NAGEL and E. M. STEIN, *On the product theory of singular integrals*, Rev. Mat. Iberoam. **20** (2004), 531–561.
[23] A. NAGEL and E. M. STEIN, *The $\bar{\partial}_b$-complex on decoupled boundarise in $\mathbb{C}^n$*, Ann. of Math. **164** (2006), 649–713.
[24] A. NAGEL, E. M. STEIN and S. WAINGER, *Balls and metrics defined by vector fields I. Basic properties*, Acta Math. **155** (1985), 103–147.
[25] J. PIPHER, *Journé's covering lemma and its extension to higher dimensions*, Duke Math. J. **53** (1986), 683–690.
[26] S. POTT and P. VILLARROYA, *A $T(1)$ theorem on product spaces*, ARXIV: 1105.2516.
[27] E. SAWYER and R. WHEEDEN, *Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces*, Amer. J. Math. **114** (1992), 813-874.

Department of Mathematics
Auburn University
Auburn, AL 36849-5310, U.S.A.
hanyong@auburn.edu

Department of Mathematics
Macquarie University
NSW, 2109, Australia
ji.li@mq.edu.au

Department of Mathematics
National Central University
Chung-Li 320, Taiwan
clin@math.ncu.edu.tw